



# VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

## FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

## ÚSTAV RADIOELEKTRONIKY

DEPARTMENT OF RADIO ELECTRONICS

## MODERNIZACE VYHODNOCENÍ ZÁZNAMU Z PALUBNÍHO REGISTRÁTORU

MODERNIZATION OF EVALUATION OF THE STORED IN-FLIGHT DATA

### BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

### AUTOR PRÁCE

AUTHOR

Roman Koníček

### VEDOUCÍ PRÁCE

SUPERVISOR

prof. Ing. Milan Sigmund, CSc.

BRNO 2019



# Bakalářská práce

bakalářský studijní obor **Elektronika a sdělovací technika**

Ústav radioelektroniky

**Student:** Roman Koníček

**ID:** 155118

**Ročník:** 3

**Akademický rok:** 2018/19

NÁZEV TÉMATU:

## Modernizace vyhodnocení záznamu z palubního registrátoru

### POKYNY PRO VYPRACOVÁNÍ:

Seznamte se s metodami záznamu a vyhodnocování letových dat a zpracujte krátkou rešerši o této oblasti v návaznosti na současný stav v Armádě ČR. Shromážděte databázi potřebných signálových dat. Naprogramujte a odzkoušejte několik metod na čištění řečových signálů pořízených analogovým magnetofonem P-503B při letu vrtulníků Mi-24V/Mi-35. Realizujte automatickou detekci hlasové komunikace v audio záznamu. Vyřešte časovou synchronizaci řečových dat s ostatními letovými daty z palubního registrátoru. Navrhněte využití vašeho modulu ve spolupráci se stávajícím systémem pro zvýšení efektivity vyšetřování leteckých incidentů.

### DOPORUČENÁ LITERATURA:

- [1] JAN, J. Číslíková filtrace, analýza a restaurace signálů. Brno: VUT v Brně, 1997.
- [2] PAŘÍZEK, J. Spolehlivost a technická diagnostika. Skriptum Univerzity obrany Brno, Brno: Univerzita obrany, 2007.

**Termín zadání:** 4.2.2019

**Termín odevzdání:** 23.5.2019

**Vedoucí práce:** prof. Ing. Milan Sigmund, CSc.

**Konzultant:** Ing. Vladan Svoboda

prof. Ing. Tomáš Kratochvíl, Ph.D.

*předseda oborové rady*

### UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

## **Abstrakt**

Práce pojednává o typech palubních registrátorů a palubních magnetofonů používaných na letecké technice Armády České republiky. Krátká rešerše pak shrnuje historii jejich vývoje v armádě. Především se zaměřuje na problematiku získávání a zpracovávání dat ze současného palubního magnetofonu P-503B. V práci jsou popisovány metody jednokanálového čištění získaného audio signálu pomocí nelineárního spektrálního odečítání a pomocí vlnkové transformace. Kvalita filtrovaných nahrávek se hodnotila subjektivní metodou. Dále je popsán detektor řeči pracující na principu měření spektrální vzdálenosti. V poslední kapitole se práce zabývá řešením časové synchronizace letových dat a hlasového záznamu. Závěrem jsou posouzeny možnosti a využití navrženého programu.

## **Klíčová slova**

Palubní registrátor, palubní magnetofon, nelineární spektrální odečítání, vlnková transformace, detektor řeči, časová synchronizace.

## **Abstract**

The thesis deals with the types of on-board registrars and on-board tape recorders used in the Air Force of the Czech Republic. Short research deals with the history of their development in the army. Above all, it focuses on acquiring and processing data from the current on-board P-503B deck. There are described the methods of single channel cleaning of acquired audio signal by nonlinear spectral subtraction and wavelet transform. The quality of filtered recordings was evaluated by subjective method. Further on, a speech detector working on the principle of spectral distance measurement is described. The last chapter deals with the solution of time synchronization of flight data and voice recording. Finally, the possibilities and use of the proposed program are assessed.

## **Keywords**

On-board registrar, on-board recorder, non-linear spectral subtraction, wavelet transform, speech detector, time synchronization.

### **Bibliografická citace:**

KONÍČEK, R. Modernizace vyhodnocení záznamu z palubního registrátoru. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2019. 53s. Vedoucí bakalářské práce byl prof. Ing. Milan Sigmund, CSc.

## Prohlášení

„Prohlašuji, že svou bakalářskou práci na téma Modernizace vyhodnocení záznamu z palubního registrátoru jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne: **20. května 2019**

.....

podpis autora

## **Poděkování**

Děkuji mému vedoucímu práce profesoru Ing. Milanu Sigmundovi, CSc za vstřícný přístup a odbornou pomoc. Dále bych chtěl poděkovat za cenné rady a trpělivost svým kolegům především Ing. Vladanu Svobodovi a Ing. Petru Střechovi. Nakonec bych chtěl nejvíce poděkovat své manželce za obrovskou podporu.

V Brně dne: **20. května 2019**

.....

podpis autora





# Obsah

1. ÚVOD .....	14
2. LETECKÉ PALUBNÍ REGISTRÁTORY .....	16
2.1 Vývoj palubních registrátorů .....	16
2.2 Vývoj palubních magnetofonů .....	18
2.3 Palubní magnetofon P-503B .....	20
2.3.1 Technická data palubního magnetofonu P-503B .....	20
2.4 Pozemní magnetofon P-504N .....	21
2.4.1 Technická data magnetofonu P-504N .....	21
3. SBĚR SIGNÁLOVÝCH DAT .....	22
3.1 Teorie akustiky a fonetiky .....	22
3.1.1 Maskovací jev .....	22
3.2 Původní získávání nahrávky .....	23
3.3 Současné pořizování nahrávek .....	24
3.3.1 Vlastnosti přenosové cesty .....	24
3.3.2 Kmitočtové charakteristiky nejdůležitějších přenosových traktů .....	25
3.3.2.1 Kontrolní bod pro nastavení příposlechu interkomu .....	25
3.3.2.2 Přenosová cesta s mikrofonom DEMŠ-1A s korekcí vyšších kmitočtů ...	26
3.3.2.3 Přenosová cesta přes mikrofón David Clark model M-7 .....	26
3.3.3 Užitečné signály na audio záznamu .....	27
3.3.4 Identifikace šumu a ruchů .....	28
4. METODY ČIŠTĚNÍ ŘEČOVÉHO SIGNÁLU .....	29
4.1 Předzpracování signálu .....	29
4.2 Použité metody čištění .....	29
4.2.1 Filtr FIR .....	29
4.2.1.1 Realizace FIR filtru v programu .....	30
4.2.2 Spektrální odečítání .....	30
4.2.2.1 Segmentace a váhování signálu v čase .....	30
4.2.2.2 Obecný popis spektrálního odečítání .....	31
4.2.2.3 Nevýhody spektrálního odečítání .....	31
4.2.2.4 Rychlá Fourierova transformace .....	31
4.2.2.5 Algoritmus nelineárního spektrálního odečítání .....	32
4.2.2.6 Nelineární spektrální odečítání s pásmovým váhováním .....	33

4.2.3	Vlnková transformace .....	34
4.2.3.1	Spojité vlnková transformace CWT .....	34
4.2.3.2	Diskrétní vlnková transformace DWT .....	35
4.2.3.3	Dyadická diskretní vlnková transformace s diskretním časem.....	36
4.2.3.4	Metody odšumování a prahování DTWT .....	36
4.2.3.5	Rekonstrukce signálu pomocí IDTWT .....	37
4.2.3.6	Funkce wavelet denoiseru v Matlabu .....	37
4.3	Porovnání metod filtrace.....	37
4.3.1	Subjektivní metoda MOS .....	38
4.3.1.1	Vyhodnocení subjektivních testů MOS .....	38
5.	DETEKCE ŘEČOVÉ AKTIVITY .....	40
5.1	Detektory řečové aktivity v algoritmech spektrálního odečítání .....	40
5.2	Zobrazení funkce použitého detektoru .....	40
6.	SYNCHRONIZACE A VYHODNOCOVÁNÍ .....	42
6.1	Typy modifikací časové osy .....	43
6.2	Popis algoritmu synchronizace .....	43
6.2.1	Detekce signalizace nebezpečné výšky radiovýškoměru.....	43
6.2.2	Lokalizace .....	44
6.2.2.1	Praktické měření změny časového měřítka na P-503B .....	44
6.2.3	Vyhledání první detekce a úprava nahrávky .....	45
6.2.4	Přiřazení detekcí.....	46
6.2.5	Úprava časové osy dat palubního registrátoru BUR-SL-1 .....	47
6.3	Vyhodnocování záznamu z palubního registrátoru .....	47
6.4	Celkové schéma kódu .....	48
7.	ZÁVĚR .....	49

# Seznam symbolů a zkratek

## Symboly:

$f_{vz}$	vzorkovací frekvence
$T_{vz}$	vzorkovací perioda
$s(t)$	signál se spojitým časem
$s[n]$	signál s diskrétním časem
$H(z)$	obrazový přenos
$z^{-n}$	Z-transformace zpoždění
$\hat{X}(\omega)$	odhad spektra signálu
$X(\omega)$	spektrum signálu
$\psi$	mateřská vlnka
$\psi^*$	komplexně sdružená mateřská vlnka
$h$	vektor hodnot impulsní charakteristiky

## Zkratky:

ADC	Analog to Digital Converter, převodník z analogového do digitálního signálu
ADRS	Aircraft Data Recording System, systém zaznamenávání letových dat
AIR	Airborne Image Recorder, zapisovač obrazu pilotního prostoru
AIRS	Airborne Image Recorder System, systém záznamu obrazu pilotního prostoru
ANN	Artificial Neural Network, umělá neuronová síť
CARE	Crash Airborne Recording Equipment, havarijní zapisovač dat
CARS	Cocpit Audio Recording System, provozní systém záznamu zvuku v pilotní kabině

CSR	Crash Survivable Recorder, z odolný zapisovač letových dat
CVR	Cocpit Voice Recorder, záznamník zvuku z pilotní kabiny
CWT	Continuous Wavelet Transform, spojitá vlnková transformace
DCT	Discrete Cosine Transform, diskretní kosinová transformace
DFT	Discrete Fourier Transform, diskretní Fourierova transformace
DLR	Data Link Recorder, zapisovač komunikace datovým spojem
DLRS	Data Link Recording System, systém záznamu komunikace datovým spojem
DTWT	Discrete Time Wavelet Transform, vlnková transformace s diskretním časem
DWT	Discrete Wavelet Transform, diskretní vlnková transformace
FDR	Flight Data Recorder, havarijní záznamník letových dat
EEPROM	Electrically Erasable Programmable Read-Only Memory, elektronicky vymazatelná paměť pouze pro čtení
FFT	Fast Fourier Transform, rychlá Fourierova transformace
EMC	Electromagnetic Compatibility, elektromagnetická kompatibilita
EMI	Electromagnetic Interference, elektromagnetické ovlivňování
FIR	Finite Impulse Response, filtr s konečnou impulsní odezvou
ICAO	International Civil Aviation Organization, Mezinárodní organizace pro civilní letectví
IDTWT	Inverse Discrete Time Wavelet Transform, inverzní vlnková transformace s diskretním časem
IFFT	Inverse Fast Fourier Transform, inverzní rychlá Fourierova transformace
GDPR	General Data Protection Regulation, Obecné nařízení o ochraně osobních údajů

GMM	Gaussian Mixture Model, model gaussovského rozložení pravděpodobnosti
HMM	Hidden Markov Model, skrytý Markovův model
MIMO	Multiple-Input Multiple-Output, vícenásobné vstup a výstupy
MOS	Mean Opinion Score, průměrná známka hodnocení
MUSHRA	Multi Stimulus test with Hidden Reference and Anchor, více-stimulový test se skrytou referencí a kotvou
NATO	North Atlantic Treaty Organization, Severoatlantická aliance
PANDA	Prostředky Analýzy Dat, softwarový nástroj pro vyhodnocení záznamu z palubních registrátorů
PESQ	Perceptual Evaluation of Speech Quality, vjemové hodnocení kvality řeči
PWM	Pulse Wide Modulation, pulzní šířková modulace
SII	Speech Intelligibility Index, index srozumitelnosti řeči
TRMS	True Root Mean Square, skutečná střední kvadratická hodnota
VAD	Voice Activity Detect, vyhodnocení detekce řeč - pausa

## Seznam obrázků

Obr. 2-1: Palubní registrátor SARPP-12GM .....	16
Obr. 2-2: Registrátor TESTER UZ-L, převzato z [3] .....	17
Obr. 2-3: Registrátor BUR-SL-1, převzato z [4] .....	17
Obr. 2-4: Palubní záznamový systém CARE, převzato z [5] .....	18
Obr. 2-5: Magnetofon MARS-BM, převzato z [6] .....	19
Obr. 2-6: Magnetofon P-507M, převzato z [7] .....	19
Obr. 2-7: Magnetofon P-503B (U3) bez krytu .....	20
Obr. 2-8: Pozemní magnetofon P-504N .....	21
Obr. 3-1: Plocha slyšení akustických signálů .....	22
Obr. 3-2: Prahy slyšitelnosti pro různé kmitočty .....	23
Obr. 3-3: Prahy slyšitelnosti pro různé úrovně akustického tlaku .....	23
Obr. 3-4: Přenosová cesta reproduktor - diktafon .....	23
Obr. 3-5: Přenosový kanál celého systému .....	24
Obr. 3-6: Odolnost nesymetrického a symetrického vedení proti impulsnímu rušení ...	25
Obr. 3-7: Přenosový kanál od generátoru .....	25
Obr. 3-8: Přenosový kanál pro mikrofón DEMŠ-1A .....	26
Obr. 4-1: Schéma použitého nelineárního odečítání .....	33
Obr. 4-2: Modifikace nelineárního spektrálního odečítání .....	34
Obr. 4-3: Blokové schéma odšumování pomocí DTWT a IDTWT .....	37
Obr. 5-1: Zobrazení VAD v audio nahrávce .....	41
Obr. 5-2: Detail úseku s VAD detekcí .....	41
Obr. 6-1: Změny časového měřítka .....	43
Obr. 6-2: Příklad detekce přítomnosti signálu nebezpečné výšky .....	44
Obr. 6-3: Úsek vysílání před první detekcí nebezpečné výšky .....	46
Obr. 6-4: Úsek vysílání uprostřed po sobě jdoucích detekcích nebezpečné výšky .....	46
Obr. 6-5: Schéma práce programu .....	48

## Seznam tabulek

Tab. 4-1: Vyhodnocení poslechových testů.....	39
Tab. 6-1: Metoda výběru detekcí z nahrávky .....	46



# 1. ÚVOD

V leteckém průmyslu je zaznamenávání, zpracovávání a archivace dat z každého letu rutinou zlepšující objektivní kontrolu nad bezpečností cestujících, posádky letounu, materiálu a letadla samotného. Letová data jsou zapisována pomocí havarijních nebo provozních letových zapisovačů. Tyto zapisované údaje, čítající i stovky zaznamenávaných parametrů současně, jsou dnes pomocí vizuálních aplikací schopny vytvořit přehledný obrázek o tom, jak se letadlo chovalo v různých letových režimech. Záznamy mohou také pomoci při výcviku pilotů a v případě závady, či havárie jsou nepostradatelná opora pro inspektory a vyšetřovatele leteckých nehod.

Letové zapisovače se začali povinně používat poprvé na dopravních letadlech v padesátých letech 20. století a během své služby prošly mnoha změnami. Tyto změny měly zásadní vliv na metody zapisování a vyhodnocování parametrů, která jsou do nich ukládána. Letové zapisovače jsou většinou ve vojenské terminologii označovány jako palubní registrátory, tímto vojenským termínem budou tedy letové zapisovače v této práci také označovány. V Armádě České republiky se provozuje většina letadel, které jsou vybaveny registrátory třetí generace (technologie s pevnou pamětí), nebo generace druhé (magnetický záznam). Nejmodernější letové zapisovače a palubní registrátory umožňují do pevné paměti zaznamenávat stovky letových parametrů současně (ADRS/FDR), provádět záznam zvuku z kabiny letadla (CARS/CVR), obraz (AIRS/AIR) nebo záznam datového spojení (DLRS/DLR) [1] [2]. Pokud dojde k incidentu v letovém provozu, jsou data z palubního registrátoru stažena do počítače s příslušným programovým vybavením. Letová data jsou dále vyhodnocována prostřednictvím programu, ve kterém je možné spustit simulaci letu a analyzovat požadovaná data. Ve stejném programu jsou také údaje uchovávány. Zvuková stopa má stejné časové měřítko jako data, a proto je možné zobrazit zapsaná data z libovolného časového úseku letu a poslouchat hlasový záznam odpovídající danému okamžiku.

U vrtulníků Mi-35 je prováděn zápis letových dat palubním zapisovačem ruské výroby BUR-SL-1 (dále jen BUR nebo registrátor). Nicméně na tento druh zapisovače není ukládán zvukový záznam z letu. Zvukový záznam je ukládán na palubní magnetofon P-503B (dále jen P-503B nebo magnetofon). Tento magnetofon provádí analogový záznam zvuku na kazetu s magnetickým drátkem. Pokud dojde na tomto typu vrtulníku k incidentu v leteckém provozu, jsou tyto záznamy vyhodnocovány odděleně, což je dosti nepraktické. Navíc analogový záznam zvuku z magnetofonu je značně nekvalitní a aplikovaná technologie nahrávání na magnetický drátek není schopna zajistit konstantní rychlost posuvu. Přehrávání záznamu z magnetického drátku na Pozemním magnetofonu P-504N vnáší do řečového signálu další šumy k záznamu a další změnu časového měřítka. Ve střednědobém plánu AČR (Armáda České republiky) je navrženo nahrazení vrtulníků Mi-24V/35 novým typem a proto se u stávajících počítá pouze se zástavbou jen nejnutnějšího avionického vybavení v rámci směrnic ICAO (International Civil Aviation Organization). Nový typ vrtulníku dosud nebyl vybrán a stroje Mi-24V/35 mají u nás v Armádě České republiky i v jiných zemích NATO (North Atlantic Treaty Organization) stále své platné místo. Proto na 22. Základně vrtulníkového letectva vznikl požadavek na modernizaci vyhodnocování záznamu z palubního registrátoru BUR-SL-1, která spočívá právě v možnosti připojení

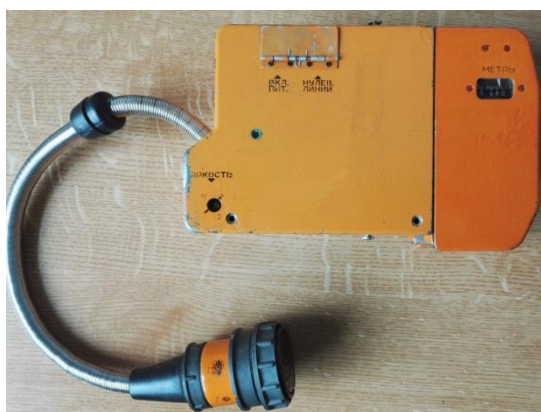
digitalizovaného záznamu a synchronizaci stopy z palubního magnetofonu P-503B k jednomu audiovizuálnímu softwarovému prostředí.

## 2. LETECKÉ PALUBNÍ REGISTRÁTORY

Letecká doprava patřila vždy mezi nejbezpečnější způsob přepravy osob a materiálu. Aby tomu tak mohlo být i nadále, musí být na bezpečnost v letecké dopravě dlouhodobě kladen velký důraz. Jak již bylo zmíněno, data ukládána do letových zapisovačů jsou po každém letu stažena a zpracována, k tomuto účelu se využívá zapisovačů označovaných jako provozní, či lehký letový zapisovač (lightweight). Při kolizi letadla jeho úkol přebírá tzv. černá skříňka, což jsou z odolnějších letové zapisovače havarijní (crash protected). [1] Z odolnějších zapisovačů musí splňovat náročnější testy pro případ nehody (normy a doporučení ICAO, např. standardy TSO-C-124 / TSO-C123a, ED-55, ED-56a), v jejich paměti (většinou typu Flash) musí zůstat informace nepoškozeny a čitelné i za těchto náročných podmínek.

### 2.1 Vývoj palubních registrátorů

Palubní registrátory první generace se začaly vyrábět v období druhé světové války a jejich vývoj dále pokračoval v 50. letech. Registrátor prováděl zápis omezeného počtu analogových signálů. Tento zápis byl prováděn mechanicky pomocí jehel na kovovou fólii. Mezi 70. A 80. lety byla vyvinuta a dále vylepšena druhá generace palubních registrátorů. Tyto registrátory již prováděly záznam signálů na plastový nebo kovový magnetický pásek pomocí magnetofonových hlav, případně pomocí světelné stopy na filmový pásek. Registrátor druhé generace SARP-12GM, který je na Obr. 2-1, využívá technologii světelné stopy na filmový pásek a je provozován již pouze na vrtulnících Mi-24V [2]. V tomto období se rovněž začaly vyvíjet palubní magnetofony pro záznam řeči. Tyto magnetofony jsou dále popsány v kapitole Vývoj palubních magnetofonů. Registrátor SARPP-12GM je možno zaměnit s minimálním zásahem do konstrukce vrtulníku výrobkem FDR-39HGM (technologie pevné paměti) firmy SPEEL Praha, s.r.o..



**Obr. 2-1: Palubní registrátor SARPP-12GM**

Na letadlech v Československu se používaly registrátory sovětské výroby typu TESTER, PARES, BUR a SARPP. Registrátor TESTER, který je na Obr. 2-2, využívá

záznamu na magnetický pásek. Registrátor BUR je již číslicový a umožňuje automatizované zpracování záznamu letu.



**Obr. 2-2: Registrátor TESTER UZ-L, převzato z [3]**

Od konce 80. let (po současnost) je vyvíjena třetí generace palubních registrátorů, ty představují plně číslicové registrátory s pevnou pamětí. Typ použité paměti je nejčastěji EEPROM (Electrically Erasable Programmable Read-Only Memory) nebo Flash [2]. Do této skupiny patří i palubní registrátory BUR-SL-1 (Obr. 2-3), které jsou instalovány do bojových vrtulníků Mi-35. Jde o plně číslicový registrátor, který má dobu přípravy k provozu do jedné minuty. Zaznamenává 16 analogových signálů rychlostí 4Hz a 16 dvouhodnotových signálů rychlostí 2Hz. Pro tuto práci bude využito synchronizace dvou dvouhodnotových signálů vyexportovaných pomocí programového vybavení PANDA do textového souboru. Délka nepřetržitého provozu je závislá na rychlosti datového toku a činí až 8 hodin při rychlosti zápisu 64bitů za sekundu. Komunikace zařízení probíhá po sběrnici RS-232. Zařízení je schopno vydržet maximální násobek přetížení až 15 g. Souprava palubního registrátoru se skládá z bloku sběru dat, paměťového zařízení a kódového zařízení [4].



**Obr. 2-3: Registrátor BUR-SL-1, převzato z [4]**

Moderní palubní registrátory třetí generace jsou vybaveny dalšími zařízeními, sloužící k záznamu zvuků a hlasu v kabině a obsahuje i nouzový maják a vysílač polohy

pro případ havárie. Hlavní představitel registrátoru třetí generace je palubní záznamový systém CARE (Crash Airborne Recording Equipment), který je zobrazen na Obr. 2-4.



**Obr. 2-4: Palubní záznamový systém CARE, převzato z [5]**

V letectvu Armády České republiky se dále využívá v různých modifikacích palubních registrátorů FDR, CSR (Crash Survivable Recorder). Všechny zmíněné typy jsou již založeny na principu zápisu do pevných pamětí (solid state). Pokud je v systému letového zapisovače zařazen i CVR, software PANDA umožňuje spolu s prohlížením parametrů letu také přehrávat audio záznam.

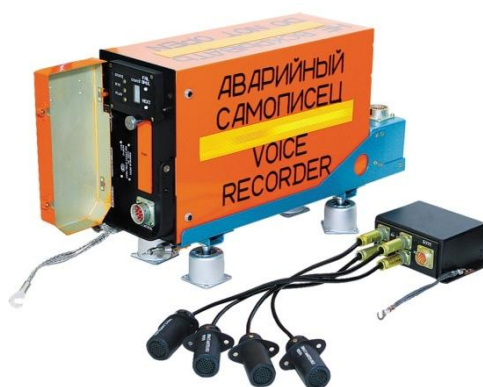
## **2.2 Vývoj palubních magnetofonů**

První palubní magnetofony zaznamenávaly zvukovou informaci na magnetický drátek nebo pásek pomocí magnetofonové hlavy.

Nevýhody prvních magnetofonů byly v omezené délce záznamu. Délka nepřetržitého záznamu se například u ruského palubního magnetofonu MARS-BM pohybovala kolem 30 minut, po té se magnetický pásek od konce přemazával. Jiný magnetofon MS-61 má délku nepřetržitého záznamu na magnetický drátek 5,5 hodiny. Nepříjemností byla nutnost převíjet drátek na začátek před každým letem, protože se jednalo o magnetofon bez reverzace. Novějším nástupcem MS-61 je magnetofon P-503B (str. 20), který má možnost automatického reverzu kazety, tudíž je záznam přemazáván jako u MARS-BM. Nevýhodou přemazávání dřívější stopy zanedbává fakt, že délka nepřetržitého záznamu je u P-503B minimálně 4 hodiny, což daleko překračuje maximální dobu letu vrtulníku Mi-24V/35, na kterém je P-503B instalován. Stávající a neméně podstatnou nevýhodou principu magnetického záznamu je znepresnění časového měřítka. Více k problematice v kapitole Synchronizace. Nekonstantní posuv pásku byl řešen připojením speciální bloku viz kapitola 2.3.1 Technická data palubního magnetofonu P-503B.



**Obr. 2-5: Magnetofon MARS-BM, převzato z [6]**



**Obr. 2-6: Magnetofon P-507M, převzato z [7]**

Později se začaly používat magnetofony, které využívaly jako paměťové médium magnetofonový pásek a informace se nahrávala opět prostřednictvím magnetofonové hlavy. Typickým příkladem je magnetofon MARS-BM, který je na Obr. 2-1. Řešení nestálosti posuvu média byla řešena obdobně, v tomto případě časovými značkami na paralelní stopu vedle zvukové nahrávky. Stějně jako třetí generace palubních registrátorů, tak i palubní magnetofony později přešly na záznam s pevnou pamětí. Současné palubní magnetofony jsou buď samostatná zařízení, která provádí záznam zvukové informace, nebo jsou integrována do palubních registrátorů. Palubní magnetofon P-507M na Obr. 2-6 je magnetofon ruské výroby, který používá pro uchování informace pevnou paměť.

V AČR jsou užívány kromě zmíněného magnetofonu P-503B magnetofony MARS-BM. Palubní magnetofon P-503B na vrtulnících Mi-24V/35, u většiny ostatních typů letadel jsou již nahrazeny zástavbou některých systémů zápisu letadel obsahující systém záznamu řeči a zvuků v kabiny (CVR (Cocpit Voice Recorder )/CARS (Cocpit Audio Recording System)). Pro srovnání vrtulníky Mi-171Š získaly díky modernizaci možnost zaznamenávat zvuk z kabiny z přílbových a ruchových mikrofónů po dobu několika hodin.



## 2.3 Palubní magnetofon P-503B

Jedná se o palubní magnetofon ruské výroby určený k jednobokovému záznamu akustického signálu na magnetický drátek ze speciální slitiny kovů.



Obr. 2-7: Magnetofon P-503B (U3) bez krytu

### 2.3.1 Technická data palubního magnetofonu P-503B

Doba spojitého záznamu	max. 9 hodin (Reverse Off) min. 4 hodiny (Reverse On)
Lineární rychlost záznamového nosiče	(100 ± 10) mm/s (výkyvy max. 7 %)
Kmitočtový rozsah zesilovače záznamu	300 – 3400 Hz (udávaná) (shodný s palubním telefonem)
Napájecí napětí	27 V ± 2,7 V
Odběr ze sítě 27V	20 W
Relativní úroveň šumu poslechu	min. 34 dB

Nosič záznamu kovový drátek typu 5.4, rozsah zaznamenaných analogových dat je od 100 Hz do 5 kHz.

Palubní magnetofon umožňuje připojení speciálního bloku reálného času, který zabezpečuje zapisování tónových pulsů s plněním (133 ± 4) Hz na nosič záznamu současně s audio záznamem. Pro získání záznamu pulsů je nutné připojit převodník k určitým pinům konektoru REZERV pozemního magnetofonu P-504N.[8] [9] Blokem reálného času ale vybavení vrtulníků nedisponuje, proto není v kapitole Synchronizace řešení dále zmiňováno. Nicméně instalací bloku reálného času by byla realizovatelná přesná synchronizace P-503B a BUR-SL-1. V kapitole Synchronizace je proto provedena pouze částečná synchronizace.

## 2.4 Pozemní magnetofon P-504N

Pozemní magnetofon P-504N na Obr. 2-1 je vybavení především pro přehrávání záznamu z palubních magnetofonů P-503B a MS-61, umožňuje také nahrávání hlasového záznamu na nosič.



Obr. 2-8: Pozemní magnetofon P-504N

### 2.4.1 Technická data magnetofonu P-504N

Lineární rychlost záznamového nosiče	(100 ± 10) mm/s ("I") (145 - 195) mm/s ("II")
Kmitočtová char. zesilovače záznamu	300 – 3400 Hz (možnost regulace útlumu až 12 dB/oktávu od 1 kHz)
Napájecí napětí	(220 ± 11) V/50 Hz nebo (115 ± 5) V/400 Hz
Odběr ze sítě	max. 75 VA
Sluchátkový výstup (TLF)	15 V / 600 Ω
Reproduktorové výstupy (GR)	2 V / 16 Ω

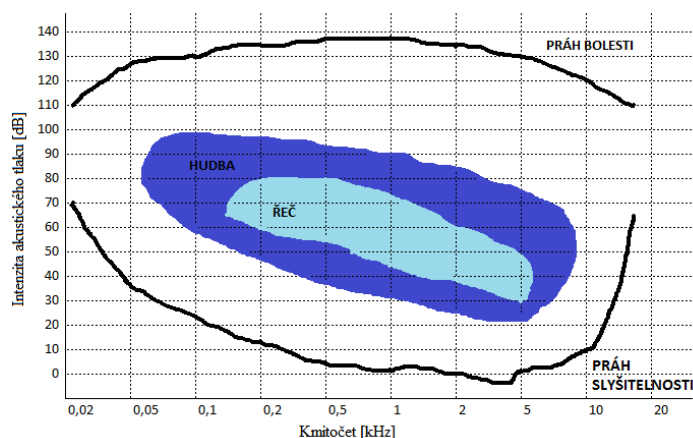


## 3. SBĚR SIGNÁLOVÝCH DAT

V této kapitole bude krátce pojednáno o základních teoretických znalosti akustiky a fonetiky a o původním a nynějším sběru signálových dat. Je zde naznačen systému přenosu zvukových signálů na pozemní části i palubě a ověřena jeho šířka přenosového pásma.

### 3.1 Teorie akustiky a fonetiky

Slyšitelné akustické signály pro lidské ucho jsou udávány pro pásmo od cca 20 Hz až do 20 kHz v úrovních od 0 až po 130 dB, jedná se o tzv. plochu slyšení (Obr. 3-1). Rozsah kmitočtového pásma řeči (fonetiky) je ale o mnoho menší. Přibližně se jedná o pásmo od 180 Hz do 6 kHz (úrovňově 30 až 80 dB). Srozumitelnosti řeči lze dosáhnout už přenosem pásma od 300 Hz až do 5 kHz. Za práh srozumitelnosti je považována hladina intenzity řeči, kdy zprávu slyšíme, ale není srozumitelná.[10]

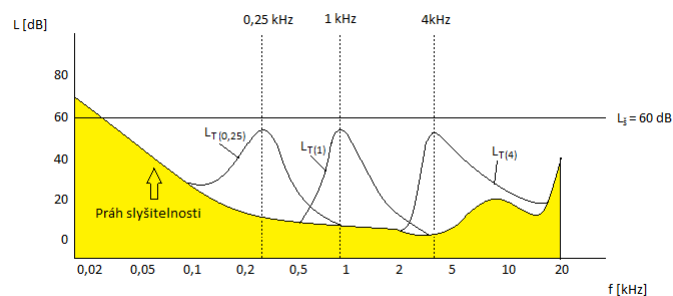


Obr. 3-1: Plocha slyšení akustických signálů

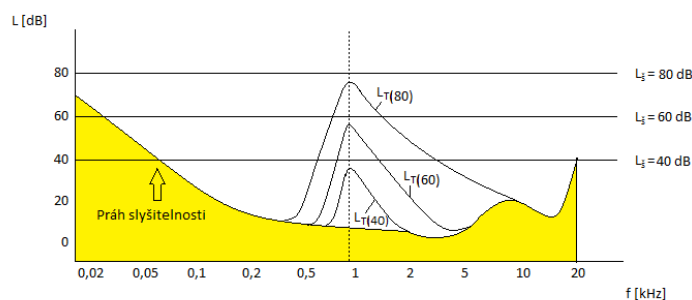
Lidské ucho je citlivé na frekvenci a intenzitu, zatímco změny fáze neregistruje.

#### 3.1.1 Maskovací jev

Dalším omezením vyplívajícím z fyziologie lidského sluchu se zabývá psychoakustika, především dochází k tzv. maskovacímu jevu slabších signálů. Maskovací jev se projevuje v kmitočtové a časové oblasti. [10]. V kmitočtové oblasti se maskovací jev projevuje mezi signály nebo tóny (trvajících déle než 200 ms) podobných frekvencí, kde intenzita jednoho signálu "přehluší" slabší signál. Signál, který úrovňově zahluší ostatní signály, upraví křivku prahu slyšitelnosti na práh současné slyšitelnosti. Signály pod touto úrovní slyšitelnosti jsou maskovány a nejsou proto pro lidské ucho slyšitelné. Křivky prahů současných slyšitelností  $L_T$  je možné vidět na Obr. 3-2 pro různé kmitočty s určitou šířkou pásma a pro různé hladiny akustického tlaku (Obr. 3-3) pro signály s určitou šířkou pásma. Závislost průběhu prahů současných slyšitelností pak závisí na kmitočtu, hladině akustického tlaku a kmitočtovém složení zvuku.



**Obr. 3-2: Prahy slyšitelnosti pro různé kmitočty**

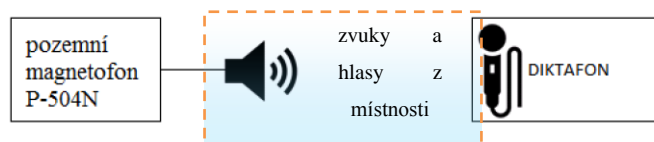


**Obr. 3-3: Prahy slyšitelnosti pro různé úrovně akustického tlaku**

Maskovací jev v časové oblasti nastane pro signály 10 ms po krátkodobém signálu vyšší hladiny. Pro delší interval maskovací jev postupně zaniká a po 200 ms zanikne zcela. Mohou být maskovány i krátkodobé impulsy 5 ms před maskujícím signálem [11].

### 3.2 Původní získávání nahrávky

Původní způsob získávání záznamu bylo nahráváním záznamu z reproduktoru P-504N do mobilního telefonu či diktafonu (Obr. 3-4).



**Obr. 3-4: Přenosová cesta reproduktor - diktafon**

Takto byly původně získávány pro další vyhodnocení všechny záznamy, jejich kvalita je ovlivněna neuspokojivou reprodukcí jak P-504N, tak jeho originálním aktivním reproduktorem. Posluchač logicky získával nahrávku s nežádoucími zvuky z místnosti.

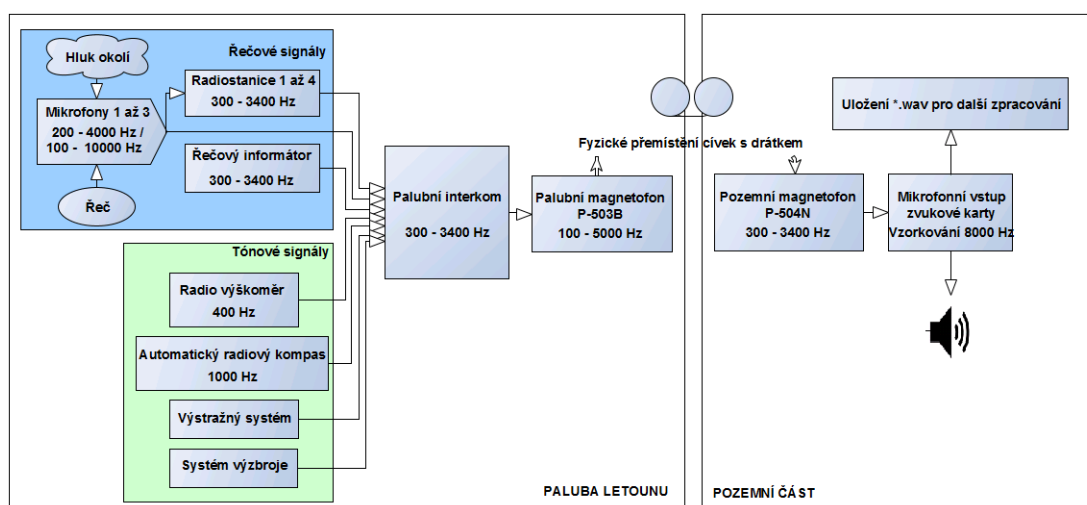
### 3.3 Současné pořizování nahrávek

Pro odstranění nevýhod akustického prostředí byl vyroben jednoduchý propojovací kabel. Signál je pomocí aplikace volně dostupného programu Audacity 2.1.0 [12] digitalizován a zobrazen.

Tímto způsobem bylo zajištěno největší množství dat a to 12 záznamů o průměrných délkách 90 minut na záznam. Vzhledem k ochraně osobních dat GDPR (General Data Protection Regulation) a utajovaných informací, které by se v nahrávkách mohli objevovat, jsou přiloženy pouze krátké audio záznamy pro ověření funkčnosti filtračních metod. Dále jsou v příloze dvě složky datových souborů, které by měly posloužit k ověření funkčnosti synchronizace radiového výškoměru a dat z registrátoru BUR-SL-1. Dle schválených žádostí od zadavatele práce, je možné použít reálná data pouze za přítomnosti zpracovatele a to za účelem ověření funkčnosti pro garantova, oponenta a komisi.

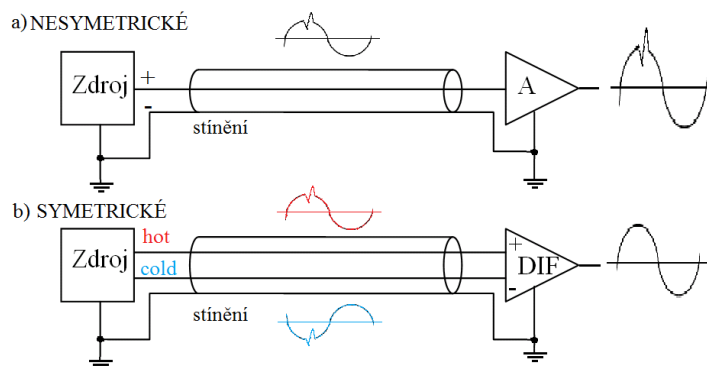
#### 3.3.1 Vlastnosti přenosové cesty

Na Obr. 3-5 jsou znázorněny přenosové cesty od všech informačních zdrojů sloučených v palubním interkomu a odtud do palubního magnetofonu P-503B. Pozemní část obsahuje pouze komponenty P-504N a zvukovou kartu osobního počítače propojenou pomocí vytvořeného kabelu.



Obr. 3-5: Přenosový kanál celého systému

Pro přenos signálů ve všech dílčích přenosových cestách na palubě je použito stíněné (Obr. 3-6) nebo nestíněné (pouze od mikrofону k připojení mikrofону) symetrické vedení, které je odolnější proti rušení EMI (Electromagnetic Interference) nežli komerčně využívané nesymetrické (Obr. 3-6). Mezi P-504N a zvukovou kartou je použito propojení symetrického a nesymetrického vedení a to z důvodu použité komerční zvukové karty, která disponuje pouze nesymetrickým mikrofonním vstupem (Jack 3.5 mm).

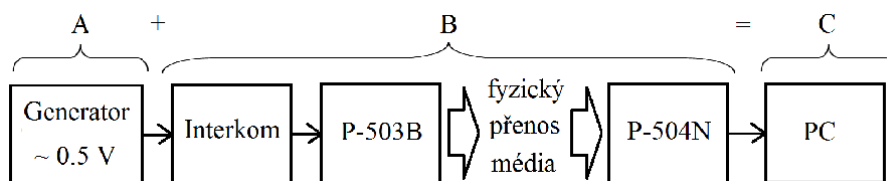


**Obr. 3-6: Odolnost nesymetrického a symetrického vedení proti impulsnímu rušení**

### 3.3.2 Kmitočtové charakteristiky nejdůležitějších přenosových traktů

#### 3.3.2.1 Kontrolní bod pro nastavení příposlechu interkomu

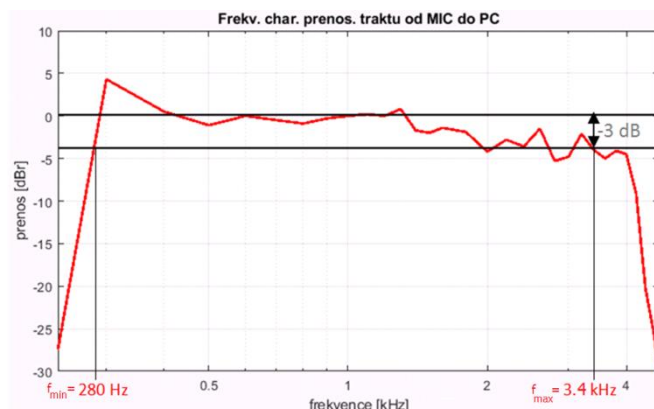
Aby bylo možné nastavit příposlech všech přilbových sluchátek zesilovače interkomu a ujednotit tak na palubě každého letounu provozní úroveň hlasitosti pro všechny členy osádky, využívá se přilbového konektoru jako kontrolního bodu. Tohoto bodu bylo využito k proměření kmitočtové charakteristiky přenosové cesty. K mikrofonnímu vstupu se připojí tónový generátor s nastavením jednotlivých frekvencí od 200 Hz do 6 kHz po určitém kroku, nastaví se konstantní velikost vstupního napětí 0.5 V TRMS (True Root Mean Square). Jelikož je nemožné změřit přenosy některých jednotlivých bloků traktu na schématu Obr. 3-7 bez zásahu do konstrukce, byl zvolen způsob měření celého přenosového traktu (na obr. označen jako B) na vrtulníku najednou. Výše popsané měření bylo proto současně nahráváno na magnetofon P-503B a záznam digitalizován. Ovládací prvky ekvalizace byly na P-504N nastaveny na nejmenší útlum vysokých a nízkých kmitočtů.



**Obr. 3-7: Přenosový kanál od generátoru**

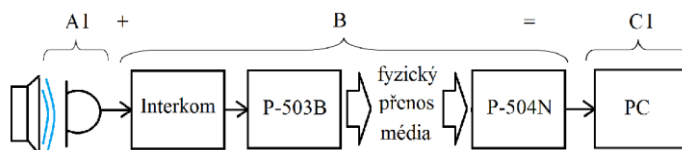
Hodnoty přenesené úrovně signálu byly získávány analýzou digitalizovaného záznamu na osobním počítači s pomocí volně dostupného nástroje Audacity 2.1.0 v oddílu Rozbor/Kreslit spektrum (nastavení: Spektrum, Hanningovo okno, velikost 4096 vzorků). Zjištěné hodnoty na Graf 3-1 kmitočtové charakteristiky byly normovány k referenčnímu bodu 1 kHz (přenos = 0 dBr) a naznačují požadavek na vzorkovací kmitočet  $f_{vz}$ .

**Graf 3-1: Frekvenční charakteristika s generátorem**



### 3.3.2.2 Přenosová cesta s mikrofonom DEMŠ-1A s korekcí vyšších kmitočtů

Mikrofon DEMŠ-1A je diferenční elektromagnetický šum potlačující mikrofon ruské výroby [13] s mikrofonním zesilovačem a udávaným rozsahem 200 - 4000 Hz. Na schématu Obr. 3-8 odpovídá bloku A1. Přenosové vlastnosti interkomu jsou označeny písmenem B, záznam z palubního magnetofonu do pozemní magnetofon a PC představuje výslednou charakteristiku Graf 3-2. Referenční hodnota charakteristiky je 1 kHz a 0.5V jako v předešlém odstavci.



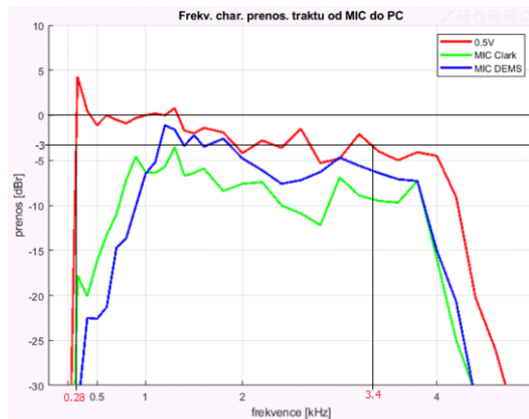
**Obr. 3-8: Přenosový kanál pro mikrofon DEMŠ-1A**

Orientačním měřením (nekalibrovaný dB - metr s udržovanou hodnotou 76 dB akustického tlaku) bylo změřeno pásmo od 200 do 6 000 Hz. Vzdálenost mikrofону od reproduktoru 20 mm. Samotný mikrofon bez zesilovače má charakteristiku kompenzující rozložení energie v lidském hlase. Rozložení energie hlasu klesá od 500 Hz výše s asymptotou 6 dB/okt., u mikrofónu naopak zisk se stejnou nebo vyšší strmostí stoupá.

### 3.3.2.3 Přenosová cesta přes mikrofon David Clark model M-7

Za stejných podmínek byl měřen i mikrofon výrobce David Clark. Jedná se o šum potlačující elektretový mikrofon s kmitočtovým rozsahem 100 až 10 000 Hz. V grafu (Graf 3-2) lze pozorovat, že kmitočtová charakteristika je velice podobná mikrofónu DEMŠ-1A. Přenosová cesta a podmínky jsou stejné jako u mikrofónu DEMŠ-1A.

**Graf 3-2: Přehled kmitočtových charakteristik od mikr. vstupu do PC**



Měřením byla poskytnuta dostatečná informace o tom, že záznam je možné vzorkovat frekvencí  $f_{vz} = 8 \text{ kHz}$ , jelikož kmitočtový rozsah záznamu je  $f_{min} = 280 \text{ Hz}$  až  $f_{max} = 3\,400 \text{ Hz}$  (pokles 3 dB od 1 kHz). Systém splňuje podmínky Shannonova-Kotělnikovova teorému:

$$f_{vz} > 2f_{max}, \quad (3-1)$$

kde  $f_{max}$  je maximální vzorkovaný kmitočet. Antialiasingový filtr odstraní kmitočty až těsně před hranicí 4 kHz a užitečné signály tak budou zachovány.

Naměřené hodnoty jsou v příloze (kód `odMICdoPC.m`).

Propustné pásmo kmitočtové charakteristiky mikrofonů DEMŠ-1A a David Clark (M-7) shora přesahuje rozsah telefonního pásma, jejich charakteristiky jsou na přenosové cestě omezeny, ale drží se jen málo od poklesu 3 dB od referenčního kmitočtu. Pro potencionálně vyšší srozumitelnost a částečné rozlišení barvy hlasu mluvčího je toto pásmo rozšířeno a po procesu filtrace použit filtr FIR na rozsah 200 až 3800 Hz 32. řádu.

### 3.3.3 Užitečné signály na audio záznamu

Užitečné informace je možné rozdělit do dvou kategorií

- řečové signály z těchto zdrojů
  - z přilbových mikrofonů posádky DEMŠ-1A (200 - 4000 Hz) nebo David Clark (M-7, 100 - 10 000 Hz),
  - rádiová korespondenci na zvolených radiostanicích 300 - 3400 Hz,
  - spouštěné namluvené fráze z řečového informátoru v rozsahu 300 - 3400 Hz,
- tónové signály
  - signál nebezpečné výšky z rádiového výškoměru (400 Hz) má nejvyšší úroveň normované energie v nahrávce a je dále využíván

pro synchronizaci, jelikož je signalizace zapsána i do parametrů BUR-SL-1,

- výstražný systému detekce raket (400 a 800 Hz),
- systému výzbroje,
- volací znaky automatického rádiového kompasu (1 kHz).

### 3.3.4 Identifikace šumu a ruchů

Spolu s užitečnou složkou signálu se mísí i aditivní níže popsané zdroje šumu:

- zdroje bílého šumu (tedy přesněji nekorelovaný náhodný šum s podobnou úrovní či výkonem v celém přenášeném pásmu) způsobený tepelným šumem součástek (přenosové kanály přijímačů a vysílačů),
- ruchem z okolí mikrofonů (při komunikaci posádky na palubě a komunikaci řídicích letového provozu je slyšitelný rozdíl šumového pozadí),
- zdroje úzkopásmového šumu,
  - motorky magnetofonů P-503B a P-504N díky 500Hz PWM (Pulse Wide Modulation) řízení (nejsilnější zdroj vícenásobného úzkopásmového šumu, při normované nahrávce dosahuje úroveň až - 27dB) a to po celou nahrávku bez významné změny spektra,
  - zdroje střídavého napětí 400 Hz - generátory, měniče,
  - zapnutá zařízení (např. klimatizace, startér - generátor),
- šum produkovaný principem elektromagnetického záznamu a přehrávání,
- pulzní šum spínání relé interkomu (objevuje se na začátku a na konci komunikace posádky a v některých případech maskuje začátek řečového signálu),
- transientní impulsní šum (především mobilní telefon, přehledový radar) se v nahrávkách objevuje díky symetrickému vedení jen ojediněle,
- výstřelový šum elektrických součástek (zanedbatelný - maskován),

## 4. METODY ČIŠTĚNÍ ŘEČOVÉHO SIGNÁLU

Metody pro zlepšení kvality řeči v zašuměném signálu mohou být rozděleny na jednokanálové a vícekanálové. K metodám s více vstupy (koherenčním) patří potlačení šumu pomocí Beam-forming, adaptivní metody potlačení šumu a systémy využívající MIMO (Multiple-Input Multiple-Output). Tato práce je omezena na použití jednokanálového čištění, kam patří zejména Wienerova filtrace, použití Kalmanova filtru, spektrální odhady a metoda založená na modelech řeči pomocí analýzy a syntézy restaurované řeči. Tato práce se bude dále zabývat jednou, ze základních metod odhadu spektra a to spektrálním odečítáním a vylepšenou Wavelet (vlnkovou) filtrací, která spadá do kategorie modelové analýzy a syntézy. [14]

### 4.1 Předzpracování signálu

Před tím, než budeme pracovat s daty digitálního signálu, je pro následné korektní zjištění jeho energie nebo počtu průchodů nulou, nutné odstranění jeho stejnosměrné složky. Nahrávky je výhodné také normalizovat. Zde je normalizováním myšlena úprava maximální úrovně stopy záznamu. Normalizování dynamického rozsahu není provedeno.

- **Ustředění**

Stejnosečná složka, tzv. DC-offset nahrávky, vzniká především při zpracování analogového signálu na digitální (vstupní obvody, operační zesilovače, chyba offsetu při kvantování v úrovni, atd.). Pakliže máme nahrávku již zajištěnou, je možné aplikovat offline metodu, přesněji jednouchý algoritmus (funkce Matlab `dcBlocker`) výpočtu střední hodnoty všech vzorků již získaného záznamu  $\bar{s}$ :

$$\bar{s} = \frac{1}{N} \sum_{n=1}^N s[n], \quad (4-1)$$

kde  $N$  je celkový počet vzorků nahrávky,  $s[n]$  představuje diskretní signál. O vypočtenou hodnotu se následně upraví celý signál.

- **Normování nahrávek**

Pro konečné doladění, je nahrávka upravena programem `normF.m` na úroveň -1 dB (cca 89%). Po filtraci spektrálním odečítáním je opět normalizována.

### 4.2 Použité metody čištění

Metody byly vybrány na základě charakteru šumu popsaného v kapitole 3, filtr FIR (Finite Impulse Response) je zařazen zejména jako pomocný prostředek pro synchronizaci dat.

#### 4.2.1 Filtr FIR

Jedná se o diskretní číslicový filtr, který se využívá k potlačení, zdůraznění či detekci (přizpůsobený filtr) frekvenčních složek v časové oblasti. FIR filtr je popisován následujícími vlastnostmi:



- **Konečná impulsní charakteristika**

Konečná impulsní charakteristika  $h = h[n]$ ,  $n \in (0, N - 1)$ , kde  $N$  je řád filtru. Diferenční rovnice vyjadřuje konečnou diskrétní konvoluci.

$$y[n] = \sum_{k=0}^{N-1} h[k] \cdot x[n - k], \quad (4-2)$$

kde  $y[n]$  je výstupní posloupnost filtru,  $x[n - k]$  jsou vstupní hodnoty a  $h[k]$  je impulsní odezva FIR filtru.

- **Stabilita**

FIR je absolutně stabilní, jestliže se všechny jeho násobné póly nachází v počátku jednotkové kružnice.

$$H(z) = \sum_{n=0}^{N-1} h_n \cdot z^{-n}, \quad (4-3)$$

kde  $H(z)$  je obrazový přenos,  $h_n$  jsou jednotlivé hodnoty impulsní charakteristiky,  $z^{-n}$  je Z-transformace pro zpoždění (např.: zpoždění pro jeden vzorek  $x[n-1]$  odpovídá výrazu  $z^{-1}X(z)$ ),  $N$  je řád filtru.

- **Linearita**

FIR filtr je lineární a proto platí princip superpozice.

- **Lineární fáze**

Může mít lineární fázovou kmitočtovou charakteristiku v celém rozsahu kmitočtů. Impulsní charakteristika musí splňovat podmínku symetričnosti nebo antisymetričnosti [15].

- **Další vlastnosti**

Filtr FIR je kauzální, odezva přichází až po podnětu, ne dříve. Dále je časově invariantní, při každém prvku posloupnosti se chová stejně.

#### 4.2.1.1 Realizace FIR filtru v programu

V programu Matlab je FIR filtr využit pro selekci kmitočtů audiosignálu pomocí pásmové propusti a to 32. řádu pro řečové pásmo 200 - 3800 Hz (`filterFIR.m`), v případě detekce signálu radiovýškoměru 370 - 410 Hz 128. řádu (`filterRV_FIR.m`).

#### 4.2.2 Spektrální odečítání

Metoda je vhodná pro čištění aditivního stacionárního (neměnné, případně pomalé změny ve spektru) šumu ve frekvenční oblasti. Výhodné je mít k dispozici části šumového signálu zcela bez řečového signálu pro co nejpřesnější odhad spektra šumu.

##### 4.2.2.1 Segmentace a váhování signálu v čase

Pro práci se signálem ve frekvenční oblasti je nutné nejprve rozložit řečový signál do mikrosegmentů, tzv. řečových úseků či rámců. Důvodem je, že samotný řečový signál

není stacionární a segmentací se z něho stává signál alespoň po částech stacionární (kvazistacionární), který potřebujeme pro extrakci stabilních parametrů řeči. Rámce jsou důležité pro výpočet rychlé Fourierovy transformace. Délka rámců je zvolena dle setrvačnosti hlasového ústrojí člověka, to je kolem 20 - 25 ms (s ohledem na zvolenou rychlost vzorkování bude pro tuto práci použito mikrosekmetů s velikostí  $N = 200$ , což odpovídá časovému rámci 25 ms). Časový signál je váhován (násoben) běžně užívaným Hammingovým oknem. V kontrastu s použitím pravoúhlého okna, ve znělých úsecích řeči nedochází při použití Hammingova okna k tak velkému útlumu kmitočtových složek s nižší intenzitou, jako při užití pravoúhlého okna. Jedná se ale o kompromis za cenu většího rozmazávání spektra [16].

Posun rámce (frame shift) je zvolen dle ověřeného 10ms posunu (centi-second vectors), v kódu uvedena délka 40 procent z hodnoty řečového rámce (Shift Percentage -  $SP = 0.4$ ). Posun rámce je zvolen tak, aby jednotlivé hodnoty sousedních rámců nebyly příliš korelovány, v některých metodách pro parametrizaci řeči se s tímto nastavením již kalkuluje (např. algoritmy obsahující HMM - Hidden Markov Model).

#### 4.2.2.2 Obecný popis spektrálního odečítání

Obecný popis spektrálního odečítání by bylo možné vyjádřit vztahem [16]

$$|\hat{X}(\omega)|^p = |Y(\omega)|^p - |\hat{N}(\omega)|^p, \quad (4-4)$$

kde  $|\hat{X}(\omega)|$  je modulové spektrum výstupního signálu z algoritmu,  $|Y(\omega)|$  je modulové spektrum vstupního signálu s šumem a  $|\hat{N}(\omega)|$  je odhad modulového spektra šumu. Indexy  $p$  pak vyjadřují mocninu ( $p = 1$  pro modul,  $p = 2$  pro výkon) spektra.

#### 4.2.2.3 Nevýhody spektrálního odečítání

Nevýhodou spektrálního odečítání jsou tzv. hudební tóny, nejzávažnějším důvodem pro jejich vznik je chybný odhad spektra šumu  $|\hat{N}(\omega)|$ . Částečným řešením je položit problematické komponenty spektra nule, tím ale zase dojde ke skokové změně hodnoty složky proti složce vedlejší, což má za následek opět vytvoření hudebních tónů v časové oblasti. Obtíže spojené se vznikem hudebních tónů bývají taktéž z důvodu špatné funkce VAD (Voice Activity Detect), použití fáze vstupního signálu pro IFFT (Inverse Fast Fourier Transform) (markantní při hodnotách  $SNR < 0$ ) [16]. Existuje mnoho variant spektrálního odečítání, každá z nich se snaží trochu jiným způsobem vypořádat s výše uvedenými hudebními tóny, dále pak s přílišným zkreslením užitečné složky nebo též s nízkou úrovní potlačení šumu. Rozdíly jsou i ve složitosti jejich implementování a ve výpočetní náročnosti. V následujících podkapitolách jsou uvedena některá používaná vylepšení spektrálního odečítání a jejich vlastnosti.

#### 4.2.2.4 Rychlá Fourierova transformace

U diskretních signálů se diskretní Fourierově řadě přiřazuje časová periodická posloupnost  $s_p[n]$  s periodou  $N$ , periodická posloupnost spektra  $S_p[k]$ , jež má periodu  $N$

$$S_p[k] = \sum_{n=0}^{N-1} s_p[n] \cdot e^{-jk\frac{2\pi}{N}n}, \quad k = 0, 1, 2, \dots, N - 1. \quad (4-5)$$

U diskrétní Fourierovy transformace jde pouze o konečnou posloupnost  $s[n]$  s délkou  $N$  spektra  $S[k]$  také o konečné délce  $N$  hodnot. Její spektrum má jednorázovou posloupnost

$$S[k] = R_N[k] \sum_{n=0}^{N-1} s[n] e^{-jk \frac{2\pi}{N} n}, k = 0, 1, 2, \dots, N - 1, \quad (4-6)$$

kde  $R_N[k]$  je váhováno koeficientem 1 pro hodnoty  $k = 0, 1, 2, \dots, N - 1$  a pro ostatní koeficientem 0. Obraz DFT (Discrete Fourier Transform) proto není periodický.

U rychlé Fourierovy transformace FFT (Fast Fourier Transform) se jedná proti DFT o zjednodušený a rychlejší výpočet díky převodům výrazu  $e^{-jk \frac{2\pi}{N} n}$  (např.:  $e^{-j \frac{3\pi}{2}} = -e^{-j \frac{\pi}{2}}$ ,  $e^{\pm j 2\pi} = 1$ , atd.). Podkapitola inspirována [17].

#### 4.2.2.5 Algoritmus nelineárního spektrálního odečítání

Než bude představena samotná metoda nelineárního spektrálního odečítání, je nutné vysvětlit vybraný algoritmus [18], který je popsán rovnicí takto

$$|\hat{S}(\omega)| = \begin{cases} |Y(\omega)| - \alpha(\omega) \cdot |\hat{N}(\omega)| & \text{jestliže } |Y(\omega)| > \beta \cdot |\hat{N}(\omega)| \\ \beta |\hat{N}(\omega)| & \text{jinak} \end{cases}, \quad (4-7)$$

kde  $|\hat{S}(\omega)|$  je odhad výstupního modulového spektra čistého signálu,  $|Y(\omega)|$  je vstupní modul spektra,  $|\hat{N}(\omega)|$  je odhad modulového spektra šumu,  $\alpha(\omega)$  je kmitočtově závislý subtrakční činitel ( $\alpha \geq 1$ ) a  $\beta$  je koeficient zbytkového šumu ( $0 < \beta \ll 1$ ).

Výhodou metody je aktualizování profilu šumu v řečových pauzách a částečná eliminace hudebních tónů. Hudební tóny lze samozřejmě upozadit přítomností vyšší hladiny původního šumu, tedy úpravou koeficientu zbytkového šumu  $\beta$  nad úroveň 0.3. Tento způsob ale degraduje účinnost odstranění šumu. Efektivní hodnota pro nastavení úrovně zbytkového šumu je  $\beta = 0.01$ . Koeficient subtrakčního činitele  $\alpha(\omega)$  je omezen nastavitelnou maximální a minimální hodnotou (nastaveny hodnoty v intervalu  $\alpha_{\min} = 1$  až  $\alpha_{\max} = 6$ ), při vysokých hodnotách činitele dochází ke značnému zkreslení a tudíž snížení srozumitelnosti na výstupu [16], např. u hodnoty  $\alpha_{\max} = 3$  jsou stále slyšitelné hudební tóny, zkreslení řeči je eliminováno. V algoritmech je hodnota  $\alpha_{\max} = 4.5$  optimální volbou, zkreslení a hudební tóny jsou neznatelné. Výpočet kmitočtově závislého odečítacího činitele  $\alpha(\omega)$  na vypočteném poměru  $SNR(\omega)$  (Signal to Noise Ratio) v logaritmické míře je uveden takto

$$\alpha(\omega) = \alpha_{\max} - \frac{\alpha_{\min} - \alpha_{\max}}{SNR_{\max} - SNR_{\min}} \cdot (SNR(\omega) - SNR_{\min}), \quad (4-8)$$

kde  $SNR(\omega)$  je vektor hodnot poměrů signálu k šumu, ostatní konstanty jsou popsány konkrétně již výše. Výpočet  $SNR$  je naznačen níže

$$SNR_i(\omega) = 20 \log_{10} \frac{|Y_i(\omega)|}{|\hat{N}_{i-1}(\omega)|}, \quad (4-9)$$

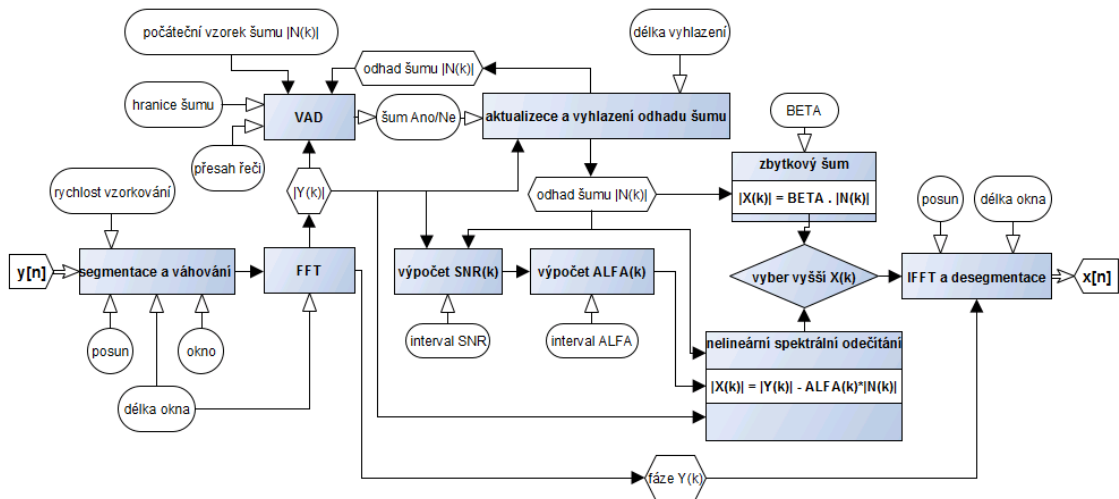
kde  $|Y_i(\omega)|$  je modul spektra vstupního signálu  $i$ -tého segmentu,  $|\hat{N}_{i-1}(\omega)|$  je odhad spektra šumu předešlého segmentu. Interval předpokládaného rozsahu  $SNR$  je taktéž přednastaven, v algoritmu konkrétně na  $SNR_{\min} = -60$  až  $SNR_{\max} = 120$  dB a jeho chybné nastavení je spjato s výskytem hudebních tónů.

Další statický parametr, kterým je markantně ovlivněna slyšitelnost hudebních tónů je tzv. smooth faktor (v kódu pod názvem `NoiseLength = 9`) pro hladký aktualizovaný odhad spektrálního profilu šumu. S vyšším faktorem `NoiseLength` ale roste četnost výskytu hudebních tónů. Aktualizování profilu šumu  $|\hat{N}_i(\omega)|$  je počítáno z předešlého odhadu spektra šumu  $|\hat{N}_{i-1}(\omega)|$  a aktuálního spektra signálu  $|Y_i(\omega)|$  následovně

$$|\hat{N}_i(\omega)| = \left( \frac{\text{NoiseLength} \cdot |\hat{N}_{i-1}(\omega)| + |Y_i(\omega)|}{\text{NoiseLength} + 1} \right). \quad (4-10)$$

Aktualizování šumu probíhá jen, když není přítomna řeč a vyhodnocení zajišťuje blok VAD. V algoritmu je použita metoda pro vyhodnocení spektrální vzdálenosti. Podrobně je použitý VAD popsán v kapitole 4 Detekce řečové aktivity.

Algoritmus vybraného spektrálního odečítání byl proti originálu kódu [18] upraven a okomentován (`nonlinear1_1.m`). Schéma na Obr. 4-1 by mělo dokreslovat princip činnosti, s tím že plné šipky znázorňují vektory rámců signálu nebo vektory parametrů.



Obr. 4-1: Schéma použitého nelineárního odečítání

#### 4.2.2.6 Nelineární spektrální odečítání s pásmovým váhováním

Pro implementaci na získaná data byla zvolena modifikace algoritmu nelineárního spektrálního odečítání [19] z předešlé kapitoly. Tato modifikace předpokládá, že šum nepostihuje řečový signál rovnoměrně a spektrum signálu je váhováno parametrem  $\delta(\omega)$ . Situace je přiblížená níže

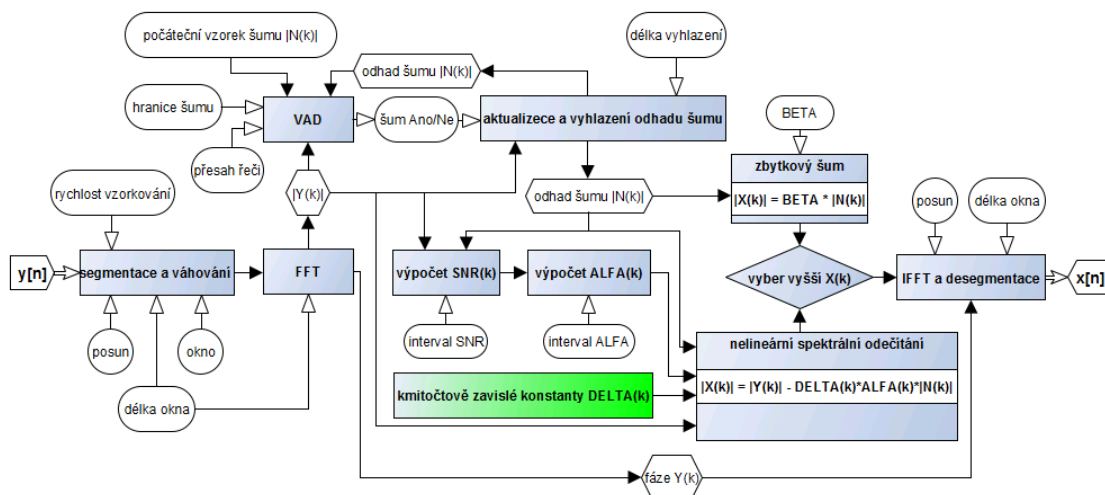
$$|\hat{S}(\omega)| = \begin{cases} |Y(\omega)| - \delta(\omega) \cdot \alpha(\omega) \cdot |\hat{N}(\omega)| & \text{jestliže } |Y(\omega)| > \beta \cdot |\hat{N}(\omega)| \\ \beta \cdot |\hat{N}(\omega)| & \text{jinak} \end{cases}, \quad (4-11)$$

kde  $|\hat{S}(\omega)|$  je odhad výstupního modulového spektra čistého signálu,  $|Y(\omega)|$  je vstupní modul spektra,  $|\hat{N}(\omega)|$  je odhad modulového spektra šumu,  $\alpha(\omega)$  je kmitočtově závislý subtrakční činitel ( $\alpha \geq 1$ ) a  $\beta$  je koeficient zbytkového šumu ( $0 < \beta \ll 1$ ), navíc je rovnice obohacena o již zmíněný koeficient vyladění  $\delta(\omega)$ . Koeficienty vyladění  $\delta(\omega)$

pro tuto konkrétní aplikaci rozdělují spektrum na tři kmitočtová pásma a přidělují jim konstantní váhy  $\delta$  následovně

$$\delta(\omega) = \begin{cases} 1.5 & 0\text{Hz} \leq \omega \leq 1\text{kHz} \\ 2.5 & 1\text{kHz} < \omega \leq 2\text{kHz} \\ 1 & 2\text{kHz} < \omega < 4\text{kHz} \end{cases} \quad (4-12)$$

Na Obr. 4-2 níže jsou označeny plnou šipkou cesty pracující s vektory hodnot.



**Obr. 4-2: Modifikace nelineárního spektrálního odečítání**

Ostatní parametry kódu (`nlinear2Multi_1.m`) jsou nastaveny stejně jako v předešlém algoritmu.

## 4.2.3 Vlnková transformace

Základním mechanismem výpočtu aplikace Matlab z Wavelet toolboxu 5.1 je vlnková transformace, konkrétně úprava spojité CWT (Continuous Wavelet Transform) a to na diskretní DWT (Discrete Wavelet Transform). Text podkapitol 1, 2, 3 a 4 je inspirován [17].

### 4.2.3.1 Spojitá vlnková transformace CWT

Obecně je vlnková transformace další metodou analýzy signálu, jako je kupříkladu FFT. Základem vlnkové transformace je mateřská vlnka  $\psi(t)$  (mother wavelet), odvozené vlnky se od této mateřské posouvají v čase a mohou se též roztahovat. Mateřská vlnka musí splňovat podmínky jako je:

- nulová střední hodnota,
- nenulová hodnota jen na konečném intervalu,
- má spojitou derivaci do řádu  $M$  vlnky (např. `symM => sym4`, pro symplet vlnu 4. ho řádu),
- existence zpětné transformace,
- ortogonalita - skalární součin mezi dvěma různými vlnkami je roven 0.

Operace s mateřskou vlnkou  $\psi(t)$  jsou popsány parametry vlnky s indexem expanze  $p$  a posuvu  $q$ , tedy  $\psi_{p,q}(t)$ . Mateřská vlnka má indexy nulové  $\psi_{0,0}(t)$ , není posunuta ani roztažena, začíná v čase  $t = 0$  a končí v čase  $t = \tau$ . Posunutím báze vlnky se změní počátek vlnky o hodnotu  $q$  na časové ose.

$$\psi_{0,q}(t) = \psi(t - q), \quad (4-13)$$

vlnka tak začíná v čase  $t = q$  a končí v čase  $t = \tau + q$ .

Expanzí mateřské vlnky  $\psi_{0,0}(t)$ , začínající v čase  $t = 0$ , a končící v čase  $t = \tau$ , vyjádřenou rovnicí níže

$$\psi_{p,0}(t) = \frac{1}{\sqrt{|p|}} \psi\left(\frac{t}{p}\right), \quad (4-14)$$

kde  $\frac{1}{\sqrt{|p|}}$  zaručuje stálou energii při změně časového měřítka. Dilatací dojde logicky ke změně spektra vlnky. Pro posunutí a roztažení současně platí vztah

$$\psi_{p,q}(t) = \frac{1}{\sqrt{|p|}} \psi\left(\frac{t-q}{p}\right). \quad (4-15)$$

Pro vytvoření spojitě funkce  $s(t)$  lze vytvořit lineární kombinaci vlnek následovně:

$$s(t) = \sum_p \sum_q b_{p,q} \cdot \psi_{p,q}(t), \quad (4-16)$$

kde koeficienty  $b_{p,q}$  udávají, jakou měrou se po vlnky  $\psi_{p,q}$  podílí na generování průběhu  $s(t)$ .

Z důvodu ortogonalit jsou všechny kombinace vlnek nulové, mimo případ kdy  $p = r$  a současně  $q = s$ . Výpočet koeficientu  $b_{p,q}$  se potom rovná:

$$b_{p,q} = \frac{\int_{-\infty}^{\infty} s(t) \cdot \psi_{r,s}^*(t) dt}{\int_{-\infty}^{\infty} |\psi_{p,q}|^2 dt}, \quad (4-17)$$

kde integrály představují normu prostoru vlnkových funkcí. Vztah pro spojitou vlnkovou transformaci je

$$S_{\text{CWT}}(p, q) = \int_{-\infty}^{\infty} s(t) \cdot \psi_{p,q}^*(t) dt. \quad (4-18)$$

#### 4.2.3.2 Diskrétní vlnková transformace DWT

Analogicky potom můžeme rovnice upravit pro dyadickou diskretní DWT. Prívlastek diskretní má transformace jenom proto, že se její parametry  $p$  a  $q$  nemohou měnit spojitě, ale s  $l$  mocninou dvou. Pro  $p = 2^l$  a  $q = i \cdot 2^l \cdot P$ , kde  $l$  a  $i$  jsou celá čísla, parametr  $P$  určuje interval skoku při posunutí, který závisí na šířce frekvenčního pásma mateřské vlnky. Čas zůstává spojitý. Takto se změní rovnice pro posun v čase a dilataci

$$\psi_{p,0}(t) = \frac{1}{\sqrt{2^l}} \psi\left(\frac{t - i \cdot 2^l \cdot P}{2^l}\right). \quad (4-19)$$

Po úpravě dostaneme rovnici diskretní dyadické vlnkové transformace

$$S_{\text{DWT}}(i, l) = \frac{1}{\sqrt{2^l}} \int_{-\infty}^{\infty} s(t) \cdot \psi^*(2^{-l} \cdot t - iP) dt. \quad (4-20)$$

Příčinou kroku dilatace vlnky o  $2^l$ -násobnou délku vlnky, dojde k odpovídajícímu zkrácení spektra vlnky na  $1/2^l$ -násobek výchozí šířky a také k posunu středního kmitočtu výchozího spektra na  $1/2^l$ -násobek původního středního kmitočtu. Každé zvýšení celého čísla  $l$  způsobí snížení kmitočtu o oktávu.

#### 4.2.3.3 Dyadická diskrétní vlnková transformace s diskrétním časem

Jak již plyne z názvu transformace, DTWT (Discrete Time Wavelet Transform) se používá pro diskretizovaný signál v čase. Pro mateřské vlnky se místo funkce vlnky použije vektor ekvidistančních hodnot. DTWT je definována následovně

$$S_{\text{DTWT}}[i, l] = \frac{1}{\sqrt{2^l}} \sum_{-\infty}^{\infty} s[n] \psi^*[2^{-l} \cdot t - iP], \quad (4-21)$$

kde  $i$  je index odpovídající úrovni (level) rozkladu nebo také koeficient dilatace,  $l$  je index, který určuje, kde se v úrovni  $i$  vyskytují koeficienty detailní a aproximační složky. Parametr  $P$  značí interval skoku při posunutí a závisí na šířce kmitočtového pásma mateřské vlnky. Detailní a aproximační koeficienty vzniknou rozkladem kmitočtového spektra na oktákové úrovně takzvanými kvadraturními frekvenčními zrcadly QMF (Quadrature Mirror Filters). Detailní koeficienty vytvoří horní propust HP a aproximační dolní propust DP. DP je po podvzorkování následně rozdělena stejně, na HP a DP, každá tato dvojice s podvzorkováním vytvoří jednu úroveň vlnkové posloupnosti. Poslední úroveň, např. úroveň 6, bude mít díky podvzorkování hodnotu vzorků o  $2^6 = 64$ krát méně než originální vstupní signál ( $T_{\text{vz6}} = T_{\text{vz}}/2^6 = 8000/64 = 125$ ).

#### 4.2.3.4 Metody odšumování a prahování DTWT

Dosud byla představena jen analýza rozkladu vstupního signálu. Pro následnou práci s koeficienty se stanoví metoda odšumování. Ta má za úkol různými způsoby zajistit znalosti o rozložení koeficientů vstupního signálu. Např. pro Wavelet toolbox Matlab se používá níže uvedených metod

- |                                                      |                                  |
|------------------------------------------------------|----------------------------------|
| - Empirical Bayes                                    | - v toolboxu pod názvem 'Bayes', |
| - Stein's Unbiased Risk Estimate                     | - 'SURE',                        |
| - Minimax Estimation                                 | - 'Minimax',                     |
| - Block James-Stein                                  | - 'BlockJS',                     |
| - False Discovery Rate                               | - 'FDR',                         |
| - Universal Threshold $\sqrt{2\ln(\text{délka}(s))}$ | - 'UniversalThreshold'.          |

Například zmíněná metoda Empirical Bayes předpokládá, že měření mají nezávislé rozložení dané předchozím modelem směsi hodnot. Metoda 'SURE' užívá pravidlo výběru prahových hodnot na Steinově nestranném odhadu rizika (funkce kvadratických ztrát), výstupem je odhad rizika pro určitou prahovou hodnotu. Podrobněji jsou metody popisovány v dokumentaci Matlab [20] ve vstupních parametrech funkce `wdenoise`.

Dalším krokem je výběr citlivosti prahování ('ThresholdMethod'). Prahování zajistí odstranění nepodstatných koeficientů širokopásmového šumu, nejvíce zkreslených složek, atd. Užívají se prahovací pravidla měkké ('Soft'), tvrdé ('Hard'), mediánové

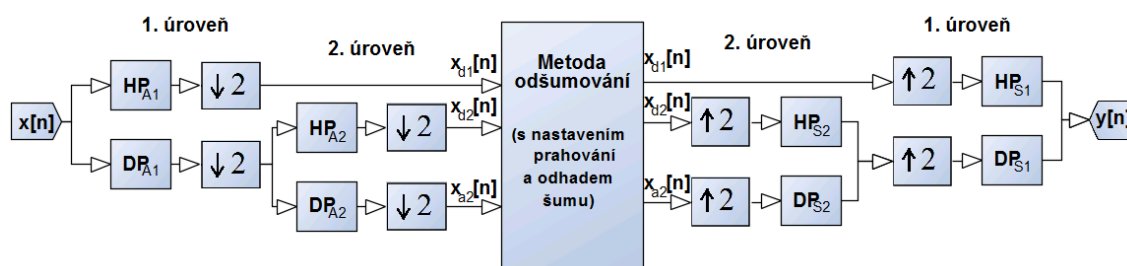
('Median'), popřípadě výpočet aritmetického průměru ('Mean'). Možnosti výběru záleží na zvolené metodě.

U většiny metod lze zvolit výpočet odhadu šumu ('NoiseEstimate') závislý ('LevelDependent') nebo nezávislý ('LevelIndependent').

#### 4.2.3.5 Rekonstrukce signálu pomocí IDTWT

Třetím a posledním stupněm pro získání vyfiltrovaného signálu je syntéza signálu pomocí IDTWT (Inverse Discrete Time Wavelet Transform). Nejdříve se nadvzorkují složky poslední (6. úrovně) aproximační a detailní složky s činitelem  $L = 2$ , dále se filtruje aproximační složka dolní propustí DP a detailní složka horní propustí HP, jedná se vlastně o vážení měřítkových funkcí  $\phi_{p,q}$  a vlnkových funkcí  $\psi_{p,q}$  s aproximačních nebo detailních koeficientů. Po filtraci se sečtou. Stejně se provede 5. úroveň, těmito iteracemi se dostaneme až k první a získáme tak výstupní filtrovaný signál.

Na Obr. 4-3 níže je blokově zobrazena dvouúrovňová rozkladová analýza (indexy A1, A2) s podvzorkováním (DTWT), blokem modelu odšumování a nadvzorkování signálů s rekonstrukční syntézou (indexy S1, S2) signálů (IDTWT).



Obr. 4-3: Blokové schéma odšumování pomocí DTWT a IDTWT

#### 4.2.3.6 Funkce wavelet denoiseru v Matlabu

Pro následné poslechové testy byl vybrán níže uvedený algoritmus potlačení šumu pomocí vlnkové transformace (obsažen v kódu GenTestWav.m).

```
out = wdenoise(in,6, ... % vstup, úroveň rozkladu
    'Wavelet', 'sym4', ... % zadání typu a řádu vlnky
    'DenoisingMethod', 'Bayes', ... % metoda (model) odšumování
    'ThresholdRule', 'Soft', ... % druh prahování
    'NoiseEstimate', 'LevelDependent'); % závislost odhadu šumu
```

### 4.3 Porovnání metod filtrace

Pro porovnání funkčnosti filtračních metod lze použít subjektivních i objektivních metod hodnocení kvality. K subjektivním metodám patří například metody MOS (Mean Opinion Score), MUSHRA (MULTI Stimulus test with Hidden Reference and Anchor), aj. K objektivním metodám patří například SNR, PESQ (Perceptual Estimation Speech Quality), SII (Speech Intelligibility Index). Pro bezchybnou funkci objektivních metod je nejvýhodnější znát, jak řečový signál degradovaný šumem, tak znát profil šumu v průběhu nahrávky nebo mít záznam čisté řeči.



### 4.3.1 Subjektivní metoda MOS

V této bakalářské práci bude kvalita nahrávky posuzována pomocí subjektivního testu MOS, jelikož je obtížné zajistit výše uvedené požadavky na správný odhad šumu nebo čistý řečový signál. Testu se účastnilo 11 osob, z toho 3 profesionálních a 8 amatérských posluchačů. Žádný s posluchačů netrpěl v době provádění poslechových testů diagnostikovanou poruchou sluchu. Všichni posluchači testu jsou povinni se účastnit v pravidelných intervalech audiometrie. Profesionálním posluchačem je označen účastník, který zná používanou terminologii komunikace, pravidelně poslouchá záznam z palubního magnetofonu a obecně rozezná případné zkreslení řeči a charakteristických signálů v nahrávce. Amatérský posluchač není seznámen s terminologií, nezná osoby na nahrávce, před tím nikdy záznamu z palubního magnetofonu neslyšel.

Hodnotící stupnice je rozdělena na pětibodovou škálu takto

- |   |                                                                               |
|---|-------------------------------------------------------------------------------|
| 5 | velké zlepšení oproti původní nahrávce                                        |
| 4 | znatelné zlepšení, ale s menší vadou                                          |
| 3 | žádné nebo minimální zlepšení / zhoršení oproti původní (referenční) nahrávce |
| 2 | znatelné zhoršení s nepříjemnou vadou                                         |
| 1 | velice nepříjemné zhoršení                                                    |

Postup a pokyny poslechových testů:

1. Kalibrace posluchače - hodnotící se nejdříve seznámí s referenční nahrávkou a následně si poslechne všechny rekonstruované nahrávky originálního vzorku za sebou.
2. Průběh hodnocení - účastník testu si poslechne nejdříve originální nahrávku, následně nahrávku rekonstruovanou. Udělí se známka. Takto si účastník testu poslechne všechny rekonstruované nahrávky. Má možnost opakovaného poslechu.
3. Posluchač přechází na další vzorek referenční nahrávky (bod 1.), takto se postup opakuje až do posledního vzorku 19.

Cílem rekonstrukčních algoritmů je především potlačení charakteristických ruchů magnetofonu, které jsou nejvíce slyšitelné v řečových pauzách. Při sepnutí spínače interkomu, či radiostanice se v tuto dobu objevuje navíc jiný, charakteristický šum, na ten není ale algoritmus zaměřen. Odfiltrování šumu při sepnutém spínači interkomu nebo rádia je výhodou. U algoritmu filtrování je samozřejmě nežádoucí slyšitelné zkreslení řeči oproti původní nahrávce.

Hodnocení poslechových testů bylo prováděno pomocí sluchátek Audio-Technica ATH-AR1iSBK (5-30 000 Hz).

#### 4.3.1.1 Vyhodnocení subjektivních testů MOS

Pro hodnocení MOS byly vybrány 3 metody filtrace a to nelineární (Nel - `nonlinear1_1.m`) a modifikované (NeM - `nonlinear2Multi_1.m`) nelineární spektrální

odečítání a odšumování pomocí vlnkové transformace (Wave - kód v kapitole 4.2.3.6). Výsledky profesionálních a amatérských posluchačů jsou v Tab. 4-1 níže. Testované záznamy zahrnují většinu druhů komunikace, která se ve všech záznamech vyskytuje. Indexace názvů vzorků je záměrná.

**Tab. 4-1: Vyhodnocení poslechových testů**

Název *.wav	Skóre profesionálních posluchačů			Skóre amatérských posluchačů			Skóre všech posluchačů		
	Nel	NeM	Wave	Nel	NeM	Wave	Nel	Nem	Wave
<b>vz02</b>	4.67	4.33	2.67	5.00	4.63	2.75	4.91	4.55	2.73
<b>vz03</b>	4.33	4.33	2.67	5.00	5.00	3.00	4.82	4.82	2.91
<b>vz04</b>	4.33	3.67	2.00	4.75	4.25	2.63	4.64	4.09	2.45
<b>vz06</b>	4.33	3.67	2.00	5.00	5.00	2.25	4.82	4.64	2.18
<b>vz07</b>	4.67	4.33	2.00	5.00	5.00	3.25	4.91	4.82	2.91
<b>vz08</b>	4.33	4.33	2.33	5.00	5.00	3.00	4.82	4.82	2.82
<b>vz09</b>	4.33	3.67	2.00	5.00	4.88	2.88	4.82	4.55	2.64
<b>vz10</b>	4.33	4.33	3.33	5.00	5.00	2.63	4.82	4.82	2.82
<b>vz11</b>	4.67	4.67	2.00	4.75	5.00	3.75	4.73	4.91	3.27
<b>vz13</b>	4.33	4.00	1.67	4.75	4.88	2.25	4.64	4.64	2.09
<b>vz15</b>	4.33	4.33	3.33	5.00	5.00	3.38	4.82	4.82	3.36
<b>vz16</b>	4.67	4.33	4.00	5.00	5.00	3.75	4.91	4.82	3.82
<b>vz20</b>	4.33	4.67	4.00	5.00	5.00	4.00	4.82	4.91	4.00
<b>vz22</b>	4.33	4.33	2.33	4.25	4.00	3.00	4.27	4.09	2.82
<b>vz23</b>	4.67	4.33	3.33	4.63	4.75	3.88	4.64	4.64	3.73
<b>vz24a</b>	4.00	3.33	1.67	4.50	3.88	2.25	4.36	3.73	2.09
<b>vz25a</b>	4.00	4.00	2.00	4.63	4.75	2.88	4.45	4.55	2.64
<b>vz26a</b>	5.00	5.00	2.00	5.00	5.00	1.88	5.00	5.00	1.91
<b>Celkem</b>	<b>4.19</b>	<b>3.98</b>	<b>2.39</b>	<b>4.59</b>	<b>4.53</b>	<b>2.81</b>	<b>4.48</b>	<b>4.38</b>	<b>2.69</b>

Dle poslechových testů profesionálních i amatérských posluchačů byl algoritmus Nel, tedy nelineární spektrální odečítání bez modifikace, hodnoceno nejlépe. Pro implementaci do celkového kódu bude proto vybrána zmíněná metoda filtrování.

## 5. DETEKCE ŘEČOVÉ AKTIVITY

Detektory řečové aktivity VAD (mají širokou škálu využití v aplikacích sloužících ke zlepšení srozumitelnosti, zvýraznění řeči, ve vokodérech, rozpoznávání řeči, ale také pro komprimaci toku dat řečového signálu (audio kodeky, vokodéry). Užívají se také k vyhodnocení začátku a konce promluvy, případně pro automatickou segmentaci promluvy.

Nejjednodušší a výpočetně nenáročné VAD vychází ze získání hodnot krátkodobé energie či vyhodnocení počtu průchodů nulou (ZRC - Zero Crossing Rate). Chybovost těchto základních metod je ale velká. Spolehlivější detektory jsou založeny na výpočtech spektrální či kepsrální vzdálenosti mezi řečí a šumem. Rozhodovací úroveň řeči může být buď pevně daná, nebo se aktualizuje dle určitých parametrů šumu (např. adaptivní práh dle dynamiky nebo charakteristiky pozadí). Jiné, stochastické detektory používají statické modely, pro správné klasifikování jsou zapotřebí trénovací data, například u GMM (Gaussian Mixture Models) nebo HMM. Trendem jsou detektory založené na umělé inteligenci ANN (Artificial Neural Network). [21]

### 5.1 Detektory řečové aktivity v algoritmech spektrálního odečítání

Pro oba popisované algoritmy spektrálního odečítání jak je popsáno v předešlé kapitole je použit detektor na bázi výpočtu spektrální vzdálenosti profilu signálu a profilu šumu. Funkce je označena pod názvem `vad.m`. Signál a šum je v každém rámci charakterizován magnitudou vypočtených komponent spektra. Profil šumu (vyhlazeného průměrováním několika po sobě jdoucích rámců šumu) v logaritmické míře je následně odečten od profilu signálu. Pakliže je odhad šumu větší než signál, spektrální vzdálenost je rovna nule.

$$\widehat{SD}(\omega) = \begin{cases} 20 \cdot (\log_{10}(S(\omega)) - \log_{10}(\widehat{N}(\omega))) > 0, \\ \text{jinak } 0 \end{cases}, \quad (5-1)$$

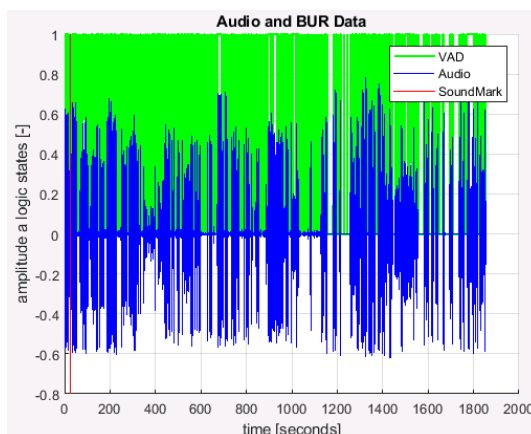
kde  $\widehat{SD}(\omega)$  je odhad modulu spektrální vzdálenosti,  $S(\omega)$  a  $\widehat{N}(\omega)$  je modul spektra vstupního signálu, resp. odhad modulu spektra šumu.

Všechny poměrné výsledné odhady  $\widehat{SD}(\omega)$  jednoho rámce jsou zprůměrovány a porovnány s fixní rozhodovací úrovní (v algoritmu `vad.m` pojmenován `NoiseMargin`, ponechán v defaultním nastavení `NoiseMargin = 3`). Jestliže výsledek převyšuje tuto mez, dojde k vyhodnocení aktivity. Při přechodu z řeči na řečové pozadí je signál aktivity ještě po několik rámců držen v této úrovni. Předpokládá se setrvačnost řeči nebo přechod do neznělých hlásek. To je ošetřeno nastavením konstanty `HangOver` v kódu (ponecháno v také v defaultním nastavení `HangOver = 8`).

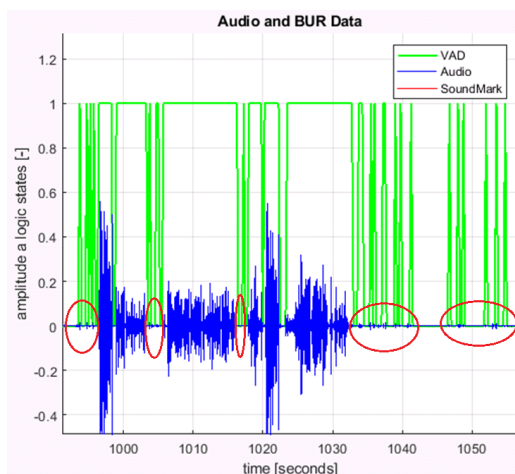
### 5.2 Zobrazení funkce použitého detektoru

Na Obr. 5-1 níže je zobrazena zelenou barvou řečová aktivita u vyfiltrovaného signálu spektrálním odečítáním s tím, že hodnoty detekcí VAD byly vyhodnoceny před filtrací.

Z obrázku je zřejmé, že nejsou detekovány jen řečové úseky, ale také méně užitečné signály na nahrávce, nicméně tyto signály měli před samotným filtrováním vyšší úroveň než vyfiltrovaný profil šumu, či ruchu magnetofonu, jsou proto správně zachovány. S tím souvisí, že práh  $\text{NoiseMargin} = 3$  je nastaven pro filtrování správně. Tento poznatek je důležitý proto, že je vhodné nahrávku reprodukovat co nejdůvěryhodněji a to pouze s potlačeným ruchem magnetofonu, nicméně samotné zobrazení detekce VAD působí rušivě.



**Obr. 5-1: Zobrazení VAD v audio nahrávce**



**Obr. 5-2: Detail úseku s VAD detekcí**

Na Obr. 5-2 je zachyceno přiblížení detailu předešlého záznamu z Obr. 5-1. Je nutno říci, že tyto signály, popsatelné jako krátké nesrozumitelné řečové úseky, mohou mít původ v nedokonalém předmagnetování (mazání v pásmu ultrazvukových vln) záznamového drátku magnetofonu, případně se také může jednat o slabý přeslech zapnutých radiostanic. Nejedná se tedy o hudební tóny vznikající užitím spektrálního odečítání, jsou slyšitelné i v originální nahrávce, na Obr. 5-2 jsou zmíněné signály zaznačeny červeně.

## 6. SYNCHRONIZACE A VYHODNOCOVÁNÍ

V tomto oddíle je řešena synchronizace získaného audio záznamu s časovou osou dvouhodnotových signálů palubního registrátoru (maximální  $f_{vz} = 4$  Hz, radiovýškoměr  $f_{vz} = 2$  Hz). Vycházíme z faktu, že máme dvě různé časové délky nahrávky a datové stopy, jednu s nesprávnou kolísající časovou osou z magnetofonu P-503B a druhou z BUR-SL-1 s časem stopy, kterou pokládáme za reálnou časovou souslednost. Rychlost posuvu drátku není konstantní z důvodu použitých mechanismů a elektronikou posuvu, která je citlivá na změny teploty [2]. Pro zajištění synchronizace musíme jednu z nich upravit tak, aby korelovala s druhou. Nabízí se dvě hlavní varianty řešení:

- **zafixovat časovou osu zvukového záznamu magnetofonu P-503B a přizpůsobit data registrátoru BUR-SL-1**

Výhody:

- Jednoduchá modifikovatelnost časové osy dat registrátoru.
- Předpokládá se jednoduchá úprava kódu v zavedeném systému PANDA.
- Nejsou nutné operace s audiosignálem, které by mohli např. při blízkých signálech výškoměru nepříjemně zkreslit řeč.

Nevýhody:

- Z pohledu pozorovatele a posluchače plyne čas na grafice v různém měřítku.
- Délka audio záznamu není stejná jako délka datové stopy BUR.
- Nutnost archivace synchronizačních časových koeficientů.

- **zafixovat data BUR-SL-1, jehož časová souslednost je považována za reálnou**

Výhody:

- Pracuje s reálným časem.
- Po převzorkování audia žádná nutnost změny algoritmu stávajícího systému vyhodnocení.
- Stejný čas audio záznamu i parametrů z palubního registrátoru.

Nevýhody:

- Výpočetní náročnost operace převzorkování, či aproximace vzorků záznamu.
- Degradace záznamu v případě chybné synchronizace.
- Opět je ale nutné zachovat původní nemodifikovanou nahrávku.

Jelikož je nezbytné originální nahrávku, tedy v původní délce a před všemi ztrátovými úpravami, archivovat ve stávajícím systému, je výhodnější upravit kód

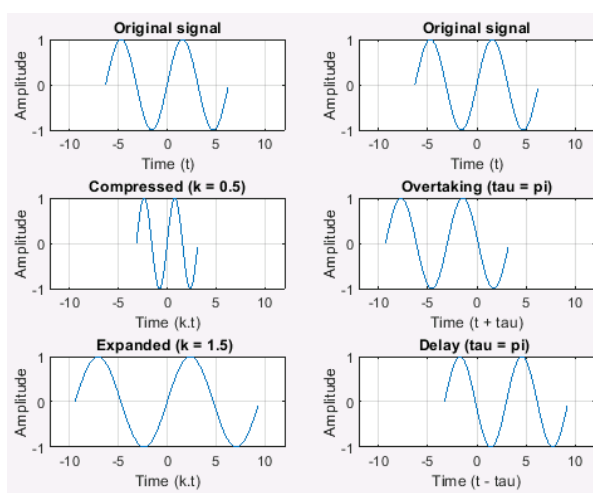
stávajícího systému PANDA, který pouze zobrazí hodnoty z palubního registrátoru na časové ose nahrávky. Případnou informaci o reálném čase z palubního registrátoru by bylo možné zobrazit dle potřeby.

## 6.1 Typy modifikací časové osy

Obecně je možné použít několik známých metod pro úpravu časové osy záznamu z registrátoru

- komprese (popřípadě kontrakce) časové osy, konstanta  $k > 1$
- expanze (popřípadě dilatace) časové osy kdy konstanta  $k < 1$
- posunutí časové osy zpožděním nebo předběhnutím

Operace jsou názorně ukázány na Obr. 6-1.



Obr. 6-1: Změny časového měřítka

## 6.2 Popis algoritmu synchronizace

### 6.2.1 Detekce signalizace nebezpečné výšky radiovýškoměru

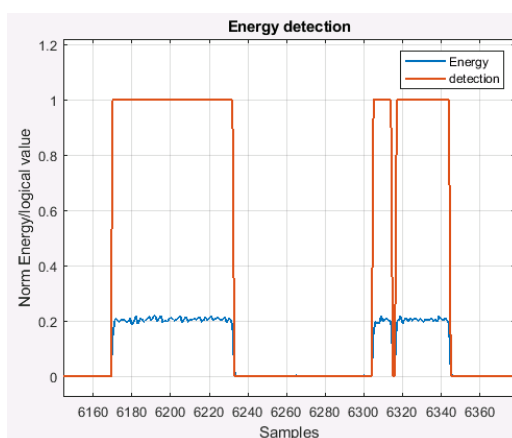
Informace o přítomnosti signálu nebezpečné výšky z palubního registrátoru BUR je obsažen v textovém souboru, který je upraven do matice pomocí algoritmu `LoadAircraftData.m`, časy detekce vypočte algoritmus `detekceRV_BUR.m` z výstupu předešlého algoritmu. Protože není svoleno ke zveřejnění výpisu ze systému PANDA, byly pro ověření vygenerovány datové soubory jednoho záznamu z algoritmu `LoadAircraftData.m` (`AirMat.mat` a `AirMatData.mat`).

Aby bylo možné získat jednotlivé časy detekcí z audio nahrávky P-503B, byly zajištěny časové detekce založené na výpočtu střední krátkodobé normované energie signálu po úsecích dle níže uvedeného vzorce

$$E = \frac{1}{l_{\text{frame}}} \sum_{n=0}^{l_{\text{frame}}-1} |s[n]|^2, \quad (6-1)$$

kde  $E$  je normovaná energie,  $l_{\text{frame}}$  je délka rámce,  $n$  je časová souslednost a  $s[n]$  analyzovaný signál.

Pro předzpracování signálu k detekci podle energie bylo nejdříve implementováno normování (`normF.m`), ustředění (`dsp.DCBlocker`) nahrávky a filtr FIR typu pásmová propust (`filterFIR_RV.m`). Algoritmus detekce `detekceRVenergy.m` rozdělí signál na rámce  $l_{\text{frame}}$  o časové délce 100 ms, rámce se nepřekrývají. Hodnoty detekce přítomnosti signálu RV jsou zapisovány po těchto úsecích, nicméně výstupem je detekce počtu náběžných hran ve zjištěných časech. Jelikož není možné z důvodů uvedených v sekci databáze poskytnout žádnou audio nahrávku, byly z kódu `detekceRVenergy.m` opět vygenerovány datové soubory jedné části záznamu, alespoň pro kontrolu funkčnosti simulace (`pocetdetekceFIR.mat` a `casdetekceFIR.mat`). Graficky je detekce na energetické bázi zobrazena na Obr. 6-2.



Obr. 6-2: Příklad detekce přítomnosti signálu nebezpečné výšky

## 6.2.2 Lokalizace

Jestliže jsou známy časy signálu nebezpečné výšky z obou zdrojů, dojde k vyhodnocení, zda se po sobě následující časy z P-503B nalézají v určitém intervalu před nebo za značkou signálu BUR. Podle množství po sobě jdoucích detekcí nalézajících se v okně intervalu je vypočten poměr a to k celkovému počtu všech značek. Algoritmus se následně skokově přesouvá od první směrem k pozdější značce detekce nebezpečné výšky BUR a opakuje vyhodnocení shody v daném intervalu. Hodnota s nejvyšším poměrem udává správné posunutí matice hodnot detekcí z nahrávky vůči parametrům z BUR. Lokalizace je řešena ve stejnojmenném souboru `lokalizace.m`.

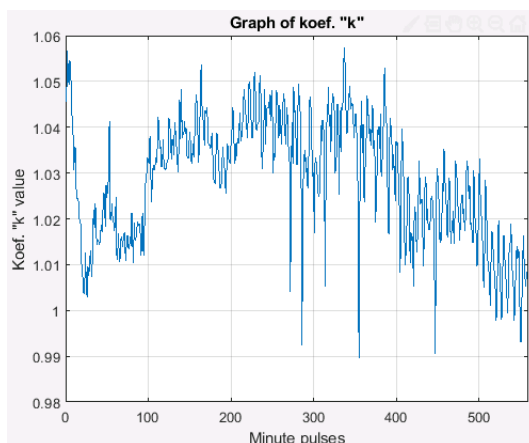
### 6.2.2.1 Praktické měření změny časového měřítka na P-503B

Pro bližší seznámení se z modifikací měřítka bylo provedeno měření jeho konstanty pomocí záznamu minutových pulsů na maximální možnou délku záznamového drátku P-503B (cca 9 hodin). Tento test byl proveden v laboratorních podmínkách, proto se z důvodů otřesů, přetížení a kolísání teplot na palubě letounu předpokládá širší interval změny koeficientů. V algoritmu `lokalizace.m` je proto nastaven podle součtu tolerancí rychlosti nahrávání a přehrávání na magnetofonech P-503B a P-504N. Jelikož každý

přístroj může mít toleranci rychlosti nahrávání, resp. přehrávání  $\pm 10\%$ , koeficient změny časového měřítka  $k$  jsou nastaveny na interval od 0.8 do 1.2.

Výsledky testu viz Graf 6-1 níže.

**Graf 6-1: Koeficienty časového měřítka**

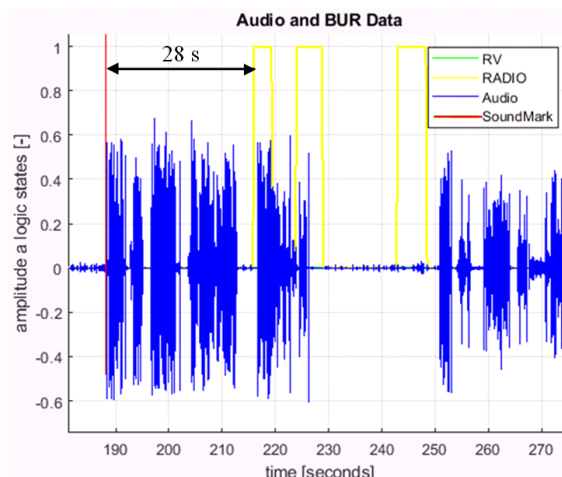


V provozu se užívá především druhá a třetí čtvrtina z celého záznamu zobrazeného na obrázku níže.

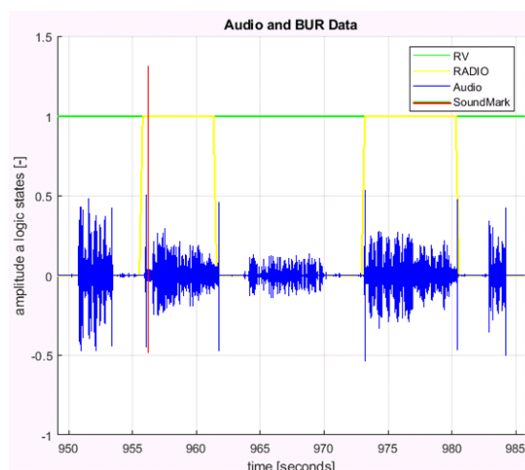
### 6.2.3 Vyhledání první detekce a úprava nahrávky

Tento krok kódu (`urcipozici.m`) vyhledá nejvyšší vypočtenou hodnotu z algoritmu lokalizace, zjistí tak nejpravděpodobnější první detekci signálu nebezpečné výšky audio záznamu, která odpovídá danému času určité detekce BUR. Úprava délky nahrávky je provedena dle trvání datového záznamu BUR, když je interval mezi začátkem a první detekcí BUR kratší než interval začátku a první detekce v nahrávce, dojde k odstranění začátku nahrávky (začátek nahrávky a dat BUR tak začínají současně v jeden okamžik). Z předpokladu, že je interval nahrávky kratší, dojde k prodloužení nahrávky (doplnění nulami) tak, aby opět data a nahrávky začínaly současně. Před první detekcí nejsou ale samozřejmě data sesouhlasena (viz Obr. 6-3), délka intervalu ořezu je násobena koeficientem  $k = 1$ . Například uprostřed intervalu (10 minut) po sobě jdoucích detekcí je odchylka neznatelná (viz Obr. 6-4).





**Obr. 6-3: Úsek vysílání před první detekcí nebezpečné výšky**



**Obr. 6-4: Úsek vysílání uprostřed po sobě jdoucích detekcích nebezpečné výšky**

Pro orientační kontrolu přesnosti synchronizace je vhodné využít signálu vysílání radiostanice (v legendě obrázků Obr. 6-3 a Obr. 6-4 označen RADIO) na grafice.

## 6.2.4 Přiřazení detekcí

Následně se vyhodnocuje platnost detekcí ze zvukové nahrávky. Podmínkou je, že se musí hledaný signál nalézat opět v určitém intervalu před nebo po signálu BUR. Jestliže se v tomto okně detekce nenachází, jsou algoritmem (`prirazeni.m`) smazány, jako je to provedeno s hodnotou 1325.3 viz Tab. 6-1.

**Tab. 6-1: Metoda výběru detekcí z nahrávky**

Vzorky detekcí z audio signálu	...	1310.7	1324.1	1325.3	1380.3	...
Detekce BUR	...	1278	1291	1344.5	1344.5	...
Výpočet koeficientu		-	13.4/13	1.2/53.5	56.2/53.5	...
Koeficient (interval 0.8 - 1.2)	...	-	1.031	0.022	1.050	...
Výběr vzorků algoritmem	...	1310.7	1324.1	-	1380.3	...

### **6.2.5 Úprava časové osy dat palubního registrátoru BUR-SL-1**

Algoritmus `upravatbur.m` pracuje s maticí dat parametrů BUR, kde jsou mimo parametrů (v tomto případě pouze dvě stopy dvouhodnotových dat) navíc přiřazeny časové značky pro každých 250 ms záznamu. Aby bylo možné správně přiřadit datovou stopu ke zvukové nahrávce, každých těchto 250 ms se přiřadí odpovídající okamžik audio stopy.

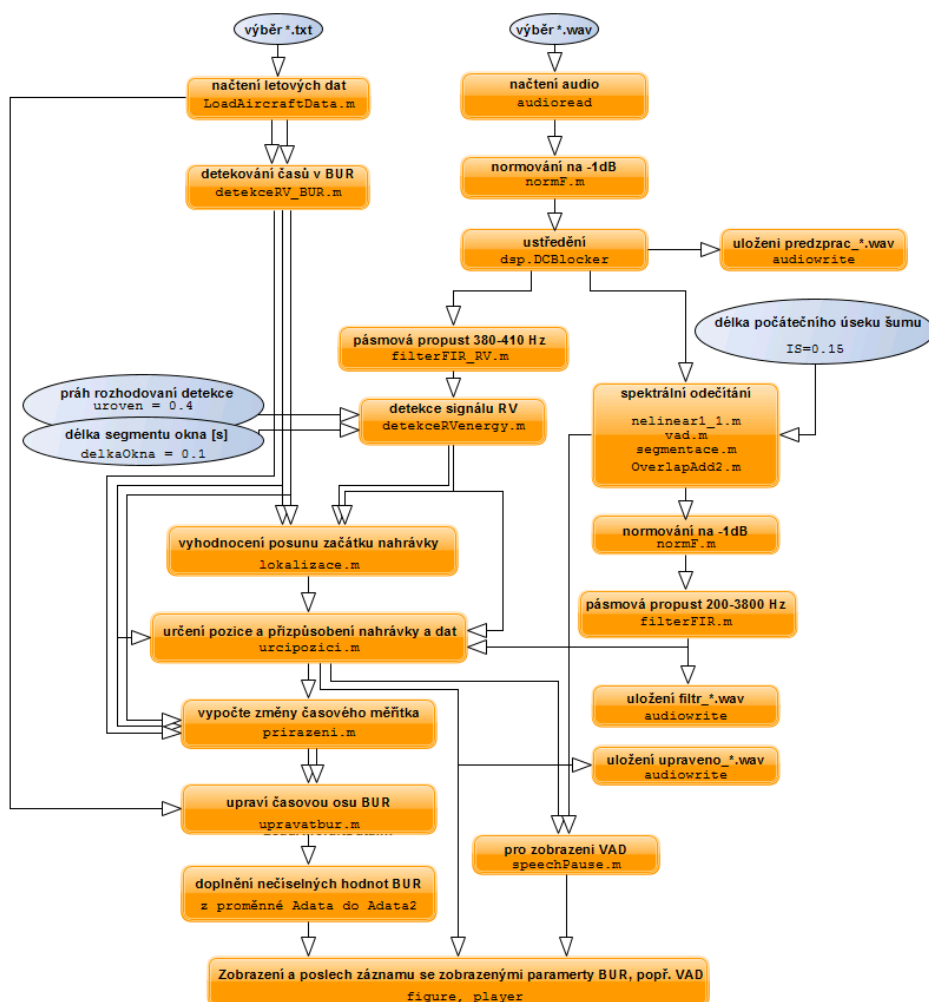
Dalším krokem bylo nahrazení nečíselných hodnot hodnotami předešlými. Následně jsou zobrazena dvouhodnotová data a nahrávka přehrána.

### **6.3 Vyhodnocování záznamu z palubního registrátoru**

Na pracovišti pro vyhodnocování záznamu se nejdříve připojí vyexportovaný textový soubor z programu PANDA (Prostředky Analýzy dat) a následně odpovídající audio soubor. Zobrazí se audio stopa se signály nebezpečné výšky a vysílání radiostanice, spustí se přehrávání filtrovaného audia. Aktuálně přehrávané místo je označeno značkovačem. Pro práci s takto zobrazenými daty je žádoucí v programu doplnit o možnost posouvání nebo skok v audio záznamu. Dle konfigurace použitého počítače je nutné snížení náročnosti zobrazení, posuvu a přiblížení či oddálení audio stopy.

## 6.4 Celkové schéma kódu

Současná podoba vyhodnocení záznamu z palubního registrátoru (Cely\_program.m) je zobrazena na schématu Obr. 6-5 níže. Hlavní funkce i všechny jednotlivé funkce Matlab kódu jsou v příloze.



Obr. 6-5: Schéma práce programu

U nejdelšího odzkoušeného záznamu (přes 2 hod. a 38 minut) trval výkon programu přes 180 sekund, z toho 85 % trvala filtrace. Test byl prováděn na počítači s konfigurací:

CPU Intel Core i7-3517U

paměť 2x 4GB DDR III

úložiště 128 GB mSATA SSD

grafická karta integrovaná NVIDIA GeForce GT 640M/2GB.

## 7. ZÁVĚR

V první kapitole je pojednáno o historii a současnosti avionických zařízení, kterými se bakalářská práce zabývá. Je zde vysvětlena problematika současného užívání tohoto vybavení na Mi-24V/35 v kontrastu s modernějším avionickým vybavením u ostatní letecké techniky. U vybavení jsou uvedeny a srovnávány některé důležité technické parametry.

Sběr signálových dat byl omezen s ohledem na povolení udělená zaměstnavatelem. V příloze jsou proto pouze krátké řečové úseky pro filtrování a vygenerované matice za běhu programu s reálnými daty. Pro sběr zvukového záznamu byla vyhotovena jednoduchá propojovací kabeláž. Vzorkování zvukového signálu kmitočtem 8 kHz bylo zvoleno v souladu se změřenou maximální užitečnou frekvencí přenosového kanálu ( $f_{\max} = 3400 \text{ Hz}$ ). Snížením vzorkovacího kmitočtu na minimální možnou hodnotu se značně snížila náročnost výpočtu algoritmů a na potřebu velikosti úložiště pro archivaci.

Pro filtraci digitalizovaného zvukového signálu byly vybrány a popsány dva druhy nelineárního spektrálního odečítání a experimentálně i odšumování pomocí vlnkové transformace. Subjektivní metodou MOS bylo vybráno nejlépe hodnocené filtrování. Jedná se o první popisovanou metodu nelineárního spektrálního odečítání. Metoda splňuje požadavek na odstranění ruchu magnetofonu a současně ponechává autentičnost celé nahrávky. Podařilo se také minimalizovat nežádoucí vlastnosti, jako je zkreslení řeči a výskyt hudebních tónů, které jsou pro spektrální odečítání charakteristické. Doporučuji nahrávku v takovéto formě poskytovat pro poslech záznamu při vyšetřování.

Aplikovaná metoda detekce řečové aktivity je použita vhodně, vycházím tak z kvalitních výsledků filtrace popsaných v předešlém odstavci. Grafické zobrazení řečové aktivity není pro vyšetřování záznamů vhodné, jejich přítomnost je spíše rušivá. Vhodnějším vylepšením by dle mého názoru bylo rozpoznávání mluvčích (zřejmě obtížně proveditelné v daném kmitočtovém rozsahu), popřípadě rozpoznávání řeči (automatický přepis komunikace by byl výhodný, za předpokladu úspěšné realizace rozpoznání mluvčího).

Vzhledem k tomu, že kvalita synchronizace je závislá přímo úměrně na počtu detekcí a rozložení detekcí signálu nebezpečné výšky, je zřejmé, že nebylo možné provést synchronizaci nahrávky a dat v každém bodě záznamů. Pro optimální synchronizaci nahrávky by byla nutná instalace bloku časových značek na vrtulník. Stávající algoritmus částečné synchronizace je pro bezchybný chod vhodné ošetřit proti nežádoucím chybám v detekci z audio záznamu paralelní podpůrnou metodou detekce. Jako podpůrná metoda synchronizace může sloužit například synchronizace s dvouhodnotovými daty vysílání radiostanice. V současnosti jsou data o vysílání použita jen jako optická kontrola úspěšnosti vyhodnocení stávajícího synchronizačního kódu.

O návrhu implementace kódu nebo části kódu do stávajícího systému pro vyhodnocení záznamů PANDA se dále jedná s prostředníkem mezi Armádou České republiky a společností SPEEL Praha s.r.o.. Nadále bude prováděno testování kódu u 22. Základny Vrtulníkového Letectva jakožto pomocné aplikace pro vyhodnocení incidentů v rámci bezpečnosti leteckého provozu. V neposlední řadě se mohou díky této

práci archivovat digitalizované audio záznamy z letu, což nebylo v minulosti kvůli chybějící technologii a vybavení možné.

# Literatura

- [1] . Dub and . Parizek, "Evolution of Flight Data Recorders", *Advances in Military Technology*, vol. 13, no. 1, p. 12, 2018.
- [2] J. Pařízek, *Spolehlivost a technická diagnostika*, Vyd. 1. Brno: Univerzita obrany, 2007.
- [3] "Polytechnické muzeum: Sbírky a fondy", *Polytechnické muzeum*, 2019. © 2005–2019 Polytechnické muzeum, Rusko, Moskva, 2019.
- [4] "NPO PRIBOR JSC: Systémy a zařízení ...", *NPO PRIBOR JSC*. Rusko, 2017.
- [5] "SPEEL PRAHA: Havarijní zapisovače CARE", *SPEEL PRAHA s.r.o.* neuvedeno, Praha, Česká Republika, 2019.
- [6] D. Yatokin, "Havárie letadla Tu-154. O černých skříňkách", in *Proza.ru*, 2016.
- [7] "Závod G.I. Petrovského: Produkty", *Závod G.I. Petrovského*. 2019.
- [8] "Magnitofon P-503B: Rukovodstvo po techničeskoj ekspluatácii". neuvedeno, Rusko, 1986.
- [9] "Magnitofon nazemnij P-504N: Rukovodstvo po techničeskoj ekspluatácii". Nižnij Novgorod, 2003.
- [10] M. Sigmund, *Rozpoznávání řečových signálů: přednášky*, Vyd. 1. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav radioelektroniky, 2007.
- [11] S. Hanus, *Základy televizní techniky III: přednášky*, Vyd. 1. V Brně: Vysoké učení technické, Fakulta elektrotechniky a komunikačních technologií, Ústav radioelektroniky, 2010.
- [12] "INSTALUJ.CZ: Tvorba hudby", *INSTALUJ.CZ*. neuvedeno, Třinec, 2019.
- [13] "Elsnabcenter: transformátory; elektromotory", *Elsnabcenter*. Elsnab, Bělorusko, 2012.
- [14] S. Vaseghi, *Advanced digital signal processing and noise reduction*, 4th ed. Chichester, U.K.: J. Wiley & Sons, 2008.
- [15] J. Jan, *Číslicová filtrace, analýza a restaurace signálů*, Vyd. 1. Brno: Vysoké učení technické, 1997.
- [16] J. Hovorka, "Algoritmy spektrálního odečítání pro zvýraznění řeči v bojových pásových vozidlech", vol. 2009, no. 35. 2009.

- [17] Z. Smékal, *Systémy a signály: 1D a 2D diskrétní a číslicové zpracování*, 1. vyd. Praha: Sdělovací technika, 2013.
- [18] Z. Esfandiar, "Berouti Spectral Subtraction", *The MathWorks, Inc.* neuvedeno, neuvedeno, 1994-2019a.
- [19] Z. Esfandiar, "Multi-band Spectral Subtraction", *The MathWorks, Inc.* neuvedeno, neuvedeno, 1994-2019b.
- [20] "Mathworks: Wavelet signal denoising", *www.mathworks.com*. The MathWorks, Inc, United States, 1994-2019.
- [21] J. Rajnoha and P. Pollák, "Detektory řečové aktivity na bázi perceptivní keprální analýzy: Konference Matlab", in *Inproceedings of MATLAB conference*, 2008.

## **Seznam příloh**

**Příloha 1 - Zdrojový kód programu, hlasové nahrávky a datové soubory jsou uloženy na přiloženém CD**